



European  
University  
Institute

DEPARTMENT  
OF POLITICAL  
AND SOCIAL  
SCIENCES

# Compensating Disadvantageous Life Events

Social Origin Differences in the Effects of Family  
and Sibling Characteristics on Educational  
Outcomes

Michael Grätz

Thesis submitted for assessment with a view to  
obtaining the degree of Doctor of Political and Social Sciences  
of the European University Institute

Florence, 19 November 2015 (defence)



European University Institute  
**Department of Political and Social Sciences**

Compensating Disadvantageous Life Events

Social Origin Differences in the Effects of Family and Sibling  
Characteristics on Educational Outcomes

Michael Grätz

Thesis submitted for assessment with a view to  
obtaining the degree of Doctor of Political and Social Sciences  
of the European University Institute

**Examining Board**

Prof. Fabrizio Bernardi, European University Institute (EUI Supervisor)

Prof. Hans-Peter Blossfeld, European University Institute

Prof. Dalton Conley, New York University

Prof. Jan O. Jonsson, Nuffield College, University of Oxford/ Swedish Institute for  
Social Research (SOFI), Stockholm University

© Grätz, 2015

No part of this thesis may be copied, reproduced or transmitted without prior  
permission of the author



## TABLE OF CONTENTS

<i>Chapter</i>	<i>Title</i>	<i>Page</i>
	Abstract	ii
	Acknowledgements	iii
1	Introduction	1
2	Making Up for an Unlucky Month of Birth in School: Causal Evidence on the Compensatory Advantage of Family Background in England	35
3	When Growing Up Without a Parent Does Not Hurt: Parental Separation and the Compensatory Effect of Social Origin	59
4	The Causal Effect of Parental Separation on Child Education: A New Instrumental Variable Approach	83
5	Competition in the Family: Estimating and Explaining Educational Inequalities between Siblings	115
6	Discussion and Conclusion	143

## ABSTRACT

This thesis is a collection of four empirical studies which analyze the effects of family and sibling characteristics on educational outcomes. The analysis in all empirical studies is guided by the compensatory effect of social origin hypothesis according to which higher social origin families can reduce the negative impact of disadvantageous characteristics and life events on their children's educational outcomes. In detail, I study the effects of month of birth, parental separation, birth order, birth spacing, and maternal age. I use data on England, Germany, and Sweden. On a methodological level, I employ natural experiments, fixed effects methods, and instrumental variable (IV) estimation in order to control for the influence of unobserved confounding variables. Overall, I find support for the initial hypothesis with respect to the effects of month of birth, parental separation, and close birth spacing. Contrary to that, I find no systematic social origin differences in the effects of birth order and maternal age on educational outcomes. In the conclusion, I discuss the implications of these findings for theories of the intergenerational transmission of education, the differences in life chances of children from socio-economically advantaged and disadvantaged families, and the allocation of resources within families. I discuss how further research could possibly test in how far differences in parental involvement between social origin groups are underlying these relationships.

**Keywords:** child development; education; inequality; family characteristics; social origin

## ACKNOWLEDGEMENTS

First of all, I have to thank Fabrizio Bernardi for the continuous support which he has provided during the four years of writing this thesis. Our conversations had a major influence on the direction of this thesis and it would have been a completely different thesis without them. Second, I would like to thank Juho Härkönen for having agreed on co-authoring a chapter of this thesis and, even more crucial, having provided me with a lot of helpful advice during the last years. Third, I would like to thank everyone at the Swedish Institute for Social Research (SOFI) and the Stockholm University Demography Unit (SUDA) where I spent four months in 2013, in particular, Erik Bihagen for making this possible. The spring term in 2014 I spent at the Department of Sociology at New York University and I would like to thank everyone whom I met there, in particular Florencia Torche and Dalton Conley who have both provided feedback on this thesis and engaged with me in some new research. Dalton Conley also agreed to serve on my thesis committee. In addition, I would like to thank Hans-Peter Blossfeld and Jan O. Jonsson who have also agreed on being part of the thesis committee and have contributed their perspectives on the content of this thesis.

Many other people have provided feedback on the thesis as a whole or on one of the four empirical chapters included in this thesis. I cannot mention them all but Kieron Barclay, Diederik Boertien, Andrés Cardona, Jani Erola, Anne-Christine Holtmann, Michael Gähler, and Berkay Özcan have provided particular critical and helpful suggestions. In addition, I would like to thank Yvonne Åberg and Helen Eriksson for sharing code which I used to construct the sample and variables used in the analysis reported in Chapter 4.

Writing this thesis would have been a lot more difficult and a lot less fun without the company and support of many people, in particular, the inequality working group at the EUI, the *fishtank* (Anna, Charlotte, Johan, and Raphael), the East Village neighbors (Anna, Vasy) and our last year visitor (Zbig), the football players, and the Tuscan hills running expert (Lorenzo).

Back home, two wonderful families have been always very keen to hear the latest progress on this work from the start till the final, final submission. Thanks for having gone through this time together with me, in particular to the one who experienced this process the closest.



## **CHAPTER 1: INTRODUCTION**

### **The Central Role of Educational Outcomes in Contemporary Societies**

In developed societies education is an important predictor of positive life outcomes, not limited to but including high occupational outcomes, high levels of well-being, good health, and longevity. The importance of education in modern societies has been predicted by modernization theory. Its proponents argued that in modern societies occupational outcomes will be fully determined by innate abilities (Bell 1973; Kerr et al. 1960; Treiman 1970). Although in contemporary societies life chances are not completely determined by innate abilities, education remains an important predictor of life chances. Therefore, the study of determinants of educational success plays an important role not only in sociology but also in disciplines bordering on sociology such as demography and economics.

There are many factors, apart from innate abilities, which are known to affect educational outcomes, including social origin (Blau and Duncan 1967; Breen and Jonsson 2005; Coleman et al. 1966; Ermisch and Francesconi 2001a; Jencks et al. 1972; Pfeffer 2008; Shavit and Blossfeld 1993) and certain family characteristics (Steelman et al. 2002). However, research focusing on the way in which family characteristics and social origin influence each other with respect to children's educational outcomes is rare. The aim of this thesis is to bring together these two strands of research by analyzing how the effects of family and sibling characteristics on educational outcomes vary by family socio-economic background. For this purpose, this thesis collects four empirical studies which analyze whether the effects of month of birth, parental separation, birth order, close birth spacing, and maternal age on child education differ with social origin.

The four separate studies test a common theoretical framework. I label this framework the compensatory effect of social origin hypothesis. The expectation derived from this framework is that the consequences of disadvantageous life events and family characteristics are less negative in higher than in lower social origin families (Bernardi 2014; Conley 2004, 2008). I expect this to be the case because higher class families have more resources available which can be mobilized in order to counteract disadvantageous life events which may negatively affect children's educational outcomes.

In the introduction to this thesis I first discuss research on educational inequalities by social origin which is followed by a discussion of research on the effects of family characteristics on educational outcomes. After that I provide the main argument bringing

together these two strands of research: the compensatory effect of social origin hypothesis. It is this theoretical framework which underlies the four empirical studies collected in this thesis. Next, I discuss the methodological issues involved in identifying the compensatory effect of social origin. I finish this introduction by providing an overview of the four empirical studies collected in this thesis.

## **The Association between Social Origin and Educational Outcomes**

An extensive literature in sociology deals with the association between social origin and educational outcomes (Breen and Jonsson 2005; DiPrete and Hout 2006). Shavit and Blossfeld (1993) argue that inequalities in educational outcomes by social origin have not been reduced in most industrial countries despite educational expansion. In recent years, more detailed research has shown that inequality in educational opportunity has actually decreased in many societies to some degree (Ballarino et al. 2009; Blossfeld et al. 2015; Breen et al. 2009; Breen et al. 2010; Jackson 2013; Tieben et al. 2010).

Independent of whether a decrease in educational inequalities can or cannot be observed, it is not contested that family socio-economic background is still considerably associated with educational outcomes. For instance, in Germany, which is one of the countries analyzed in this thesis, the probability to attend the upper track in secondary school (*Gymnasium*) at age 16 to 17 years is about 46 percentage points higher for children from families in which at least one parent successfully graduated from the upper track compared to children from families with a lower level of parental education (the figure is based on my own calculations based on the SOEP data employed in Chapter 3 and Chapter 5).

Whereas descriptive evidence on the association between different measures of social origin and educational outcomes is available in abundance, there is less research that tests the mechanisms underlying educational inequalities. There are many theories aimed at explaining how social origin influences educational outcomes but empirical tests of these mechanisms are rare, in particular against each other in the same study and employing a research design which controls for unobserved heterogeneity. For that reason, we still do not know much about why educational inequalities by social origin emerge and why they are so persistent across countries and cohorts.

Based on previous research, we can distinguish at least four mechanisms that are assumed to produce the effect of social origin on educational outcomes. First, a part of the association between family background and educational outcomes may be mediated by the

transmission of genes from parents to their children (Asbury and Plomin 2014). Research in behavioral genetics argues that large parts of education are heritable (Plomin et al. 2013). This high heritability of education is not only due to the heritability of intelligence but also to the heritability of many other psychological traits such as self-efficacy, personality, or behavioral problems which correlate with education (Krapohl et al. 2014). Outside of behavioral genetics it is, however, debated how reliable the estimates of heritability are as they are constructed comparing correlations in educational outcomes between identical and non-identical twins (Goldberger 1979; Jencks 1980; Manski 2013). For this reason the role genes play as a mechanism underlying educational inequalities is, to a large extent, unclear. Recent evidence based on genetic risk scores suggests that about 1/6 of the association between maternal and offspring education is genetically transmitted—although this may be a lower bound estimate which is likely to change upwards with improvements in the measurement of genetic risk scores of educational attainment (Conley et al. 2015; Rietveld et al. 2013).

Second, a popular approach among researchers in the field of social stratification argues that educational inequalities by social origin are a product of rational decision making (Becker 2003; Boudon 1974; Breen and Goldthorpe 1997; Breen et al. 2014; Davies et al. 2002; Erikson and Jonsson 1996; Gambetta 1987; Goldthorpe 2007; Morgan 2005). According to these theories, children and their families make educational decisions based on the perceived benefits, costs, and success probabilities of educational transitions. The Breen-Goldthorpe model assumes that parents make educational decisions so that their children achieve at least the same occupational position as they have (Breen and Goldthorpe 1997). This motive leads children from higher social origin families to pursue longer educational careers in order to minimize the risk of downward social mobility. The same motive leads children from lower social origin families to choose shorter educational pathways because they can achieve their parental class position by means of shorter educational pathways. Empirical evidence on the key mechanism underlying this model is mixed. The studies which arguably provide the most stringent test of the main underlying mechanism according to the Breen-Goldthorpe model, the motivation to avoid downward social mobility, fail to find empirical evidence in favor of it (Gabay-Egozi et al. 2010; Stocké 2007). What is more, I am not aware of any study testing the rational choice mechanisms employing a research design which controls for omitted variable bias.

Third, educational aspirations and motivations also play a central role in the Wisconsin model of status attainment (Sewell et al. 1969; Sewell and Hauser 1975, 1980). In this research tradition, educational aspirations are a result of peer influences and reactions to

previous academic performance. Since pupils are segregated into schools, educational inequalities persist through these peer influences. As a result of these peer influences children from higher social origin families have higher educational aspirations. With respect to other theoretical backgrounds, other authors have also pointed to the role of peers in influencing educational and occupational decisions (Akerlof 1997; Manzo 2013). Newer empirical research by Bozick et al. (2010) supports to some degree the view that academic performance shapes educational aspirations through the application of individual-fixed effects models but they also show that for a large proportion of students, especially those from higher social origin families, educational aspirations are stable throughout their school careers.

Fourth, the cultural capital hypothesis argues that educational positions are reproduced through the transmission of cultural practices and knowledge from one generation to another (Bourdieu 1984; Bourdieu and Passeron 1977; De Graaf et al. 2000; DiMaggio 1982; Jæger 2011; Lareau and Weininger 2003; Willis 1977). According to Bourdieu (1984) a certain combination of cultural competences, called *habitus*, is transmitted from parents to their children, largely through processes of which they are not conscious. The *habitus* of the upper class makes educational success more likely. Complementary to what Bourdieu argues for the upper class, Willis (1977) describes in an ethnographic study of British adolescents a working-class culture which devalues educational success. DiMaggio (1982) argues that it is precisely for children from low social origin families that cultural capital matters most because high levels of cultural capital enhance upward mobility. The cultural capital hypothesis is still influential in contemporary sociological work, in particular in ethnographic studies, such as Lareau's (2011) *Unequal Childhoods*. According to Lareau (2011), parents from different social classes raise their children using very distinctive parenting styles which are consequential for their children's life chances. Using a combination of family- and individual-fixed effects models, Jæger (2011) provides quantitative evidence that cultural capital indeed affects children's test scores.

This thesis does not test these four competing mechanisms against each other. Contrary to that, I argue that there is a fifth mechanism which can explain the high level of educational inequalities observed in contemporary industrialized societies. According to the compensatory effect of social origin hypothesis, higher social origin families can compensate for the negative consequences of disadvantageous life events or characteristics (Bernardi 2014; Conley 2004, 2008). The impact that disadvantageous life events and characteristics have on educational outcomes has probably become more important over time and many of these traits which negatively influence future life chances are concentrated in lower social origin families

(McLanahan 2004). Both these processes may contribute to educational inequalities by social origin. I will expand on this argument in the section on the compensatory effect of social origin.

According to the compensatory effect hypothesis, focusing on educational inequalities by social origin, without taking into account how family and sibling characteristics influence educational outcomes, does not present an adequate picture of how educational inequalities emerge. The effects of family characteristics and social origin manifest themselves simultaneously. These interactions should be taken into account in order to provide a complete picture of inequalities in educational outcomes and of the mechanisms that bring them about. But before discussing how these two research fields can be brought together via the compensatory effect of social origin hypothesis, I first introduce central results of research on the effects of family and sibling characteristics on educational outcomes.

### **Family and Sibling Characteristics and Inequalities in Educational Outcomes**

A considerable amount of research has dealt with the associations between certain family and sibling characteristics and educational outcomes (Steelman et al. 2002). More recent research has applied causal identifications strategies in order to test whether the associations between family characteristics and child education are due to underlying causal effects or whether selection processes bring about the observed associations. This literature has led to new insights so that many of the conclusions reported in the review article by Steelman et al. (2002) have to be updated given the results of this newer research.

Most prominently, a huge number of studies has argued that a larger family size is associated with lower educational outcomes (Blau and Duncan 1967; Blake 1985, 1989; Downey 1995; Hanushek 1992). More recent research, however, suggests that family size has no causal effect on educational outcomes but that the effects of family size on child education are dominated by those of birth order (Black et al. 2005; Conley and Glauber 2006; Guo and VanWey 1999). An instrumental variable (IV) approach by Jaeger (2008) leads to different results. However, he uses birth spacing as an IV which is probably violating the exclusion restriction of having no direct effect on educational outcomes since other research argues that close birth spacing leads to lower educational outcomes (Buckles and Munnich 2012; Pettersson-Lidbom and Skogman Thoursie 2009).

Other sibling and family characteristics, however, probably have a causal impact on educational outcomes. Examples of such characteristics include month of birth, parental

separation, birth order, birth spacing, and parental age. This thesis focuses on analyzing the variation in the causal effects of these family and sibling characteristics on educational outcomes.

Previous research has shown that each of these characteristics affects child education. Month of birth affects educational outcomes via the school entry age in countries with strict school entry admission rules (Bedard and Dhuey 2006; Bernardi 2014). Parental separation has been shown to be negatively associated with educational outcomes (Amato 2010; Jonsson and Gähler 1997). More recent evidence suggests that part of this association is brought about through a causal effect (McLanahan et al. 2013). Recent research, mainly focusing on differences within families, has shown that higher birth order children achieve lower educational outcomes (Black et al. 2005; Härkönen 2014). Close birth spacing is negatively associated with educational outcomes (Powell and Steelman 1993). Causal identification strategies have been employed to estimate the effect of birth spacing on child education with the result that at least part of this association is causal (Buckles and Munnich 2012; Pettersson-Lidbom and Skogman Thoursie 2009). It has been shown that a higher parental age leads to improved educational outcomes, both in cross-sectional data (Mare and Tzeng 1989) and in a research design which use only variation within families (Kalmijn and Kraaykamp 2005).

Most of the research on the effects of family and sibling characteristics on educational outcomes did not analyze how the effects of these characteristics differ with social origin. If there is heterogeneity by social origin, these studies misrepresent the true effects of these characteristics on educational outcomes. There are good reasons to believe that this is the case as I will discuss in the next section which develops the compensatory effect of social origin hypothesis.

### **The Compensatory Effect of Social Origin**

There are good theoretical reasons to expect effects of family and sibling characteristics on educational outcomes to differ with social origin. I summarize these theoretical expectations under the framework of the compensatory effect of social origin hypothesis because they predict that the consequences of disadvantageous characteristics are less negative for children from higher social origin families. Furthermore, there is already empirical evidence that the effects of some characteristics vary by family socio-economic background in the way predicted by this hypothesis.

On the one hand, for theoretical reasons a compensatory effect can be expected because advantaged families have more resources available to them in order to counteract the effects of disadvantageous life events and characteristics (Bernardi 2014). Conley (2004, 2008) develops the argument that siblings in higher social origin families are more similar in their educational and occupational outcomes than siblings in lower social origin families. This is because higher social origin families have more resources available in order to compensate for disadvantageous life events to ensure that their children are less affected by them.

This perspective also echoes the above-mentioned work by Lareau (2011) on different childrearing practices in lower and upper class families. These different styles of raising children may become particularly important in those cases in which children and families are confronted with disadvantageous life events which endanger the future educational and occupational outcomes of their children.

The compensatory effect of social origin hypothesis suggests that children from lower and higher social origin families follow “diverging destinies” (McLanahan 2004) not only with respect to the occurrence of disadvantageous life events but also that the negative consequences of disadvantageous life events are more pronounced in lower social origin families. If this hypothesis can be confirmed with respect to many influencing mechanisms, the compensatory effect may actually be a more important mechanism for the reproduction of educational inequalities than cumulative advantage (DiPrete and Eirich 2006).

On the other hand, there is empirical evidence already available that the effects of some life events vary in the direction expected by the compensatory social origin hypothesis, mainly through research on the consequences of early child health on later life outcomes (for an overview see Almond and Mazumder 2013). For instance, Almond et al. (2009) show that the negative effects of random exposure to higher radiation in utero due to the Chernobyl nuclear accident on child cognitive performance are concentrated in families in which the father has a low level of education. Similar results have been obtained for the effects of birth weight (Conley and Bennett 2001; Torche and Echevarría 2011). Bernardi (2014) adds to this literature by showing that the negative effects of school entry age, an event which is not connected to birth endowments, on grade retention in France are concentrated in low-SES families. Ermisch and Francesconi (2013) show that maternal employment during childhood has stronger negative effects on child education for children with low educated mothers than for children with highly educated mothers.

Certainly, these life events differ in their characteristics and they may also differ in the mechanisms which result in the observed compensatory effect. Parents may also respond

differently to different sorts of disadvantage (Almond and Mazumder 2013). Nevertheless, a compensatory effect of social origin has been observed for these diverse life events. Hence, it is interesting to analyze whether a compensatory effect can be observed in further instances. This will also help to develop the underlying theory of the compensatory effect.

In the previous section I argued that the compensatory effect of social origin can contribute to educational inequalities by social origin. This is the case for two reasons. First, lower class children are more often affected by these disadvantageous life events (McLanahan 2004). If these life events were indeed having negative causal effects, on educational outcomes, they imply lower levels of educational attainment simply because these life events happen more often in lower class families.

Second, the compensatory effect of social origin hypothesis argues that the consequences of these life events are more negative for children from lower class families. For that reason, the occurrence of a disadvantageous life event or characteristic may reduce the educational outcomes of a lower class child more than the educational outcomes of an upper class child. As a result, the average gap in educational outcomes between lower and upper class children is larger than in the counterfactual situation in which the life event would not exist, e.g. in a world in which there were no parental separations.

These implications are restricted to factors which influence educational outcomes independent of how families allocate their resources within families. In the case of birth order and differences in birth spacing this implication may not hold because lower class families may adapt their investment strategies and invest more in the child in the more advantageous position which may result in the end in an unchanged average educational attainment of the family. Hence, in the latter case within-family inequality does probably not change average class differences in educational attainment.

The central aim of this thesis is to test to which life events and family characteristics the compensatory effect framework can be fruitfully applied and to test in which instances a compensatory effect can be observed. It is against this theoretical background that I study how the effects of month of birth, parental separation, birth order, close birth spacing, and parental age on educational outcomes differ by family socio-economic background.

The compensatory effect of social origin hypothesis argues that the causal effect of family and sibling characteristics differs with social origin. Differences in associations between family characteristics and educational outcomes across social origin groups could be a result of differential selections into family characteristics by social origin. For this reason, research designs that do not help to support causal claims are insufficient as tools to study the

compensatory effect of social origin hypothesis. It is essential to test the theory by applying research designs that allow me to interpret effect estimates as representing causal effects. Therefore I discuss the issue of causal inference in the next part of this introduction.

## **Identification of Causal Effects**

Each chapter of this thesis aims at estimating the causal effect of at least one life event or characteristic (month of birth, parental separation, birth order, birth spacing, and maternal age) on child education. I place particular emphasis on the question whether these effects vary with social origin. This aim of the thesis requires the application of research designs which allow me to interpret the obtained estimates of events and characteristics as the causal effects of these events and characteristics on children's educational outcomes.

This thesis is concerned with what has been labelled "effects of causes" (Goldthorpe 2001, 2007, in press) or "forward causal questions" (Gelman and Imbens 2013). This follows the logic that first it has to be found out whether an effect does indeed underlie an association before the causes will be studied. In my conclusion I will discuss how future research may test mechanisms bringing about the causal effects and the additional challenges involved in such an enterprise.

Theory is important for the identification of causal effects, largely because theory is needed to explain why certain variables affect other variables (Esser 1996). Goldthorpe (2001, 2007, in press) argues that mechanisms which underlie the relation between these variables have to be proposed in order to identify causal effects. These mechanisms postulate possible ways in which actors may bring about a causal effect. In the context of the effects studied in this thesis, the actors are mostly children, their parents, and their siblings.

I agree that mechanisms and theory play a role in identifying a causal effect. With respect to the causal effects analyzed in this thesis, there are hypothesized mechanisms that bring about the effect of these characteristics on child education. In each chapter I discuss in detail the underlying mechanisms of the causal effects under study. With respect to the effects of month of birth on educational outcomes, the underlying mechanisms are the maturity effect, the relative age effect (Bedard and Dhuey 2006), and the differences in the length of time children spend in reception class. With respect to the effects of parental separation on child education, it has been argued that the underlying mechanisms are changes in parental involvement and in economic circumstances as a consequence of parental separation as well as stress experienced by the child during the separation process (Amato 2010; Jonsson and

Gähler 1997). Zajonc and Markus (1975) argue that birth order and birth spacing influence child cognitive development through changes in the intellectual family environment with each additional family member. A further mechanism could be that the time parents can spend with each child is decreasing with each new family member (Price 2008). Hence, each younger sibling spends less time with the parents as his siblings did at each age of childhood (if parents were distributing their time equally between their children). Close birth spacing may lead to additional time pressure on the parents (Powell and Steelman 1993). Maternal age may be influencing children's educational outcomes because mothers develop more skills during the course of their lives and this may positively affect their parenting and, through parenting, their children's educational outcomes (Kalmijn and Kraaykamp 2005).

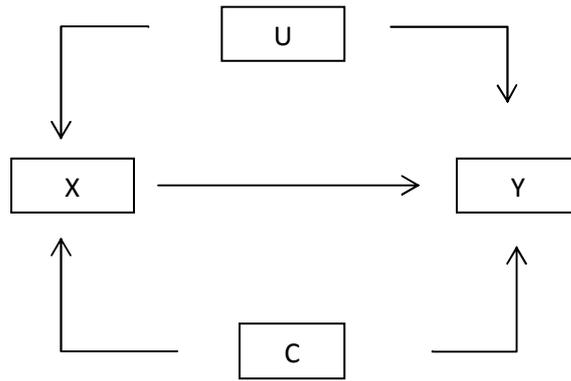
While I agree that mechanisms play an important role and always should be hypothesized as underlying causal effects, I follow those researchers who argue that postulating mechanisms is not enough in order to identify causal effects since unobserved variables may lead to selection into the family and sibling characteristics under consideration and this may bias naive estimates (Almond and Mazumder 2013; Angrist and Pischke 2009, 2010; Firebaugh 2008; Moffit 2005; Morgan and Winship 2015; Winship and Morgan 1999). "Theory has to be anchored in reality. Sooner or later, invariance needs empirical demonstration, which is easier said than done." (Freedman 2004: 277)

The challenge presented by selection on unobserved variables can be illustrated with a simple example. Figure 1.1 represents a graphical representation of the identification problem, that is of estimating the causal effect of a variable  $X$  on a variable  $Y$  when an unobserved variable  $U$  is influencing both  $X$  and  $Y$  (Freedman 2004).

$C$  is observed and can therefore be controlled for. Hence, standard regression or matching methods, as they are usually applied in sociology, can take into account selection on these observed variables. However, the more important threat to identification of the causal effect of  $X$  on  $Y$  comes from the unobserved variable  $U$ . In order to identify the causal effect of  $X$  on  $Y$ , research has either to control for  $U$  by means of employing a research design which allows researchers to control for unobserved variables or to demonstrate that there is no unobserved variable  $U$  influencing the relation between  $X$  and  $Y$  (Moffit 2005).

It is nearly impossible to demonstrate that there is no unobserved variable  $U$  influencing the relation between  $X$  and  $Y$ . Therefore I apply in this thesis design-based approaches which control for selection into family characteristics on unobserved variables (Almond and Mazumder 2013; Angrist and Pischke 2009, 2010; Firebaugh 2008; Moffit 2005; Morgan and Winship 2015; Winship and Morgan 1999).

**Figure 1.1** A graphical representation of the identification problem



This approach to causal analysis defines a causal effect based on the comparison of alternative states of  $X$ , the observed state of  $X$  and the one of an unobserved counterfactual situation (Morgan and Winship 2015). Since it is not possible to observe the counterfactual situation at the individual level, causal effects can only be estimated at the aggregated group but not at the individual level. The most general causal effect, the average treatment effect  $\delta$ , is then the average of the difference between the expected value of one state  $Y^1$  and the expected value of a counterfactual state  $Y^0$ :  $E[\delta] = E[Y^1 - Y^0] = E[Y^1] - E[Y^0]$  (Morgan and Winship 2015).

I put particular emphasis on clarifying the assumptions which allow researchers to interpret estimates as causal effects. Since the counterfactual outcome can never be observed all empirical analyses are based on assumptions in order to identify causal effects (Moffitt 2005). However, some assumptions are more likely to be met than others and scientific evidence based on weaker assumptions is less likely to be refuted by further research. If studies identifying causal effects under different assumptions reach similar conclusions, this should increase confidence in the results of the analyses.

It is important to note that the counterfactual approach is an important departure from “variable sociology” (Esser 1996) and the tradition of causal statements based on statistical regularities of “robust dependence” (Goldthorpe 2001). Design-based approaches do not disregard the importance of theory in the identification of causal effects. “Design-based studies are distinguished by their prima facie credibility and by the attention investigators devote to making both an institutional and a data-driven case for causality.” (Angrist and Pischke 2010: 5) Theory is needed to justify the assumptions underlying any research design. For that reason, missing theory is actually not a critique which should be raised against the research design-based approach to causal inference. To my mind, this approach is very much

in line with the following conclusion by Goldthorpe: “In sum, the argument is, again as with Freedman, that establishing causation cannot result from statistical procedures alone but must be dependent upon some subject-matter theoretical input relating to how the data under analysis are produced.” (Goldthorpe in press: 126)

The challenge, then, is to find research designs which allow researchers to interpret estimates as causal effects under the weakest possible assumptions (Manski 1995). All empirical chapters collected in this thesis are concerned with estimating causal effects in situations in which conditioning is ineffective because selection is on unobserved variables (Morgan and Winship 2015). I view this as a rather conservative approach given the fundamental nature of the problem of causal inference (Angrist and Pischke 2009).

Any conditioning approach such as matching or regression analyses requires the complete model to be specified in order to yield unbiased and consistent estimates of causal effects. This is a highly unrealistic assumption in most applications but, in particular, as I argue in each chapter, in the situations analyzed in this thesis. Without controlling for unobserved heterogeneity, the postulation of theoretical mechanisms alone does not allow researchers to argue that an effect is indeed causal because unobserved variables may always confound the relationship.

I employ natural experiments (Chapter 2), family-fixed effects models (Chapter 3 and 5), and Instrumental Variable (IV) estimation (Chapter 4). In each chapter I will discuss the method employed and the assumptions under which the estimates provided by these methods can be interpreted as causal effects. Therefore I will refrain from a discussion of these methods in this introduction. Previous research has employed all these methods, although, as I will show in the next section, they are not often employed in sociological research on educational inequalities.

### **Causal Analysis in Sociological Research on Educational Inequalities**

Causal identification strategies do not seem to be widely employed in sociological research on educational inequalities. In order to analyze systematically whether this is the case, I conducted a short survey of leading journals publishing work on educational inequalities by sociologists. For this purpose I analyzed all papers dealing with educational inequalities which were published in *American Journal of Sociology*, *American Sociological Review*, *Sociology of Education*, *Social Forces*, and *Research in Social Stratification and Mobility* between 2010 and 2014. The purpose of this survey was to figure out how many published

articles have employed a method which controls for the influence of unobserved confounding variables. Table 1.1 summarizes the main findings of this short survey.

**Table 1.1** Articles on educational inequalities employing causal identification strategies which control for unobserved heterogeneity in major sociological journals (2010-2014)

<b>Journal</b>	<b>Number of Articles on Educational Inequalities</b>	<b>Number of Articles on Educational Inequalities Using a Quantitative Method</b>	<b>Articles on Educational Inequalities Employing a Causal Identification Strategy</b>
<i>American Journal of Sociology</i>	6	5	Lauen and Gaddis (2013)
<i>American Sociological Review</i>	13	9	Hasan and Bagde (2013) Legewie and DiPrete (2012)
<i>Social Forces</i>	20	20	Andrew (2014) Bozick et al. (2010) Marteleteo and de Souza (2013)
<i>Sociology of Education</i>	45	43	Bernardi (2014) Burdick-Will (2013) Caudillo and Torche (2014) Grigg (2012) Hanselman et al. (2014) Jæger (2011) Kurlaender and Grodsky (2013)
<i>Research in Social Stratification and Mobility</i>	36	36	Karlson (2011)

*Source:* Own calculations.

One article applying an IV approach was published in *Research in Social Stratification and Mobility* between 2010 and 2014 (Karlson 2011). One article published in *American Journal of Sociology* used individual-fixed effects models (Lauen and Gaddis 2013). One article published in *Social Forces* employed family-fixed effects models (Andrew 2014), one used individual-fixed effects models (Bozick et al. 2010), and one used a natural experiment (Marteleteo and de Souza 2013). Two papers published in *American Sociological Review* employed a natural experiment (Hasan and Bagde 2013; Legewie and DiPrete 2012). With respect to *Sociology of Education*, seven articles published between 2010 and 2014 employed research designs which control for unobserved heterogeneity. Grigg (2012) employs individual-fixed effects models. Jæger (2011) combines family- and individual-fixed effects models to analyze the effect of parental cultural capital on educational outcomes. Kurlaender and Grodsky (2013) employ a natural experiment through a change in admission policy at the

University of California. Burdick-Will (2013) as well as Caudillo and Torche (2014) use several fixed effects methods, at the school and the local level, to isolate the causal effect of exposure to local violence on child education. Hanselman et al. (2014) employ a randomized field experiment. Finally, Bernardi (2014) employs the same natural experiment as Chapter 2 of this thesis.

To sum up this literature review, few studies in sociological research on educational inequalities employ causal identification strategies. One could argue that through the choice of analyzed journals I have somehow reduced the possibility of finding causal identification studies because sociologists may prefer to publish such studies in more interdisciplinary journals, such as *Demography* or *Journal of Human Resources*, where such studies are more commonly published. But if this were the case (which I have not tested but I believe it is), this only leads to the follow-up question why sociologists do not publish these studies in journals which are more sociologically oriented.

My survey results receive further support through a study conducted in Germany. In a survey of studies using panel data in the two leading German sociological journals, Giesselmann and Windzio (2014) show that only a minority of studies which use panel data employs individual-fixed effects models which would allow researchers a causal interpretation of effect estimates. Methods which do not control for unobserved heterogeneity are far more dominant in published research using panel data.

All in all it seems that the scarce use of design-based approaches to causal inference in sociology is something which requires an explanation. One argument could be that sociologists are more interested in description than in explanation. Many influential sociological studies focus mainly on description. For instance, research on social mobility is largely dominated by studies on social mobility trends over time. However, such a position is unsatisfactory since there is also an interest in explaining these trends. Indeed, sociologists have developed many theories which propose explanations of social phenomena, for example, the theories I mentioned above on how educational inequalities by social origin emerge. Putting these theories to critical tests requires, in my view, causal identification strategies because the mechanisms are supposed to be causal in nature. For that reason it is surprising that few sociological papers employ these techniques.

In this respect, it is also striking that there are many papers in the journals I analyzed that use causal language in the abstract or title of the paper. Although these papers talk about “effect”, “impact”, and “influence”, most of them do not employ a research design that would allow researchers to identify a causal effect.

In a review article, Gangl (2010) argues that the inevitable trade-off between internal and external validity is one of the reasons for the scarce employment of causal identification research designs in sociology. In contrast to other disciplines, such as psychology and political science, sociology may be particularly concerned with external validity. This focus on external validity may be the reason why few sociologists employ causal identification research designs. Although this objection is right, I would argue that the gains in internal validity outweigh the losses in external validity. Furthermore, one way to achieve external validity may be through several studies employing different causal identification strategies in different countries, cohorts, and contexts. The unity of these studies may then allow researchers to judge the external validity of causal claims.

Even more reason for the objection of sociologists to causal identification methods, according to Gangl (2010), is the fact that central sociological concepts such as class, gender, and race cannot be manipulated. With respect to this point, it is, however, important to note that in this thesis I do not aim at estimating the causal effect of family background but at estimating how the causal effects of certain chosen family and sibling characteristics vary by family background.

In a recent essay, Watts (2015) argues that much of sociological thinking is, without acknowledging it, heavily influenced by common sense and how results can be understood intuitively. One could then argue that one of the reasons for the slow adoption of causal inference methods in sociology may be that these methods possibly lead to results which may be less intuitively understandable. To give an example: when comparing results from a research design which controls for omitted variable bias to the naive estimates, the question emerges how the differences between the naive and the causal estimate can be explained. In other words, the question is what unobserved variables cause selection as to certain characteristics. This question cannot be answered empirically; one can only speculate about possible factors that play a role but the factors cannot be tested precisely because they are unobserved. I guess this is a feature of causal identification that seems unsatisfactory to many sociologists. However, to my mind therein lies precisely the beauty of these methods: that they do not require the variables for which they control to be measured or to be even known. This makes these methods universally applicable. In addition, measurement error is less likely to bias results in studies that use these research designs than in studies which rely on conditioning on observed variables.

## **Countries Included in the Analysis**

The primary aim of the four empirical studies collected in this thesis is to provide estimates of the variation of causal effects of life events and characteristics by family socio-economic background. The primary rationale for the selection of cases was the availability of data which would allow me to conduct such an enterprise. In particular, the chosen research designs made it necessary to search for appropriate data that are not available for many countries. I will now explain my selection of countries used in the specific studies.

The first study of the thesis uses data on England because of the strict admission rules, based on month of birth, in the English education system. Furthermore, the rich data sources, provided in particular through the linkage of survey data with administrative data on school grades, allow us to measure educational performance at different ages. For these reasons, England is an ideal case to study the evolution of month of birth effects, and their variation with social origin, over the school career. I have chosen to conduct this study using specifically English data after searching data from many countries and analyzing their educational systems.

The second study uses data on Germany from the German Socio-Economic Panel Study (SOEP) because this is one of the longest-running and largest panel studies worldwide. Thanks to its structure as a household panel that collects information on all children growing up in the household, these data can be used to apply family-fixed effects models. A further advantage of the SOEP data is that it has different measures of important educational outcomes (see next section). The sample sizes of the family-fixed effects models is larger than the one of many previous studies which applied this approach to study the effects of parental separation in other countries (Ermisch and Francesconi 2001b; Sandefur and Wells 1999).

The third study applies an instrumental variable (IV) approach to Swedish registry data. This large administrative data source makes it possible to avoid the problem of the instrument being too weak to estimate the causal effect of parental separation.

The fourth study again uses data on a sample of siblings derived from the SOEP. This is one of the largest survey data sets allowing researchers to analyze within-family inequality in educational outcomes. The sample size of these models is large compared to other research using survey data (Conley and Glauber 2008; Conley et al. 2007).

Given that this thesis uses data on three different countries, it is legitimate to ask whether differences between countries influence the results of the specific studies. However, such a question cannot be answered without conducting similar analyses in different

countries. Given that, with the exception of the second and the third study, the studies analyze the impact of different characteristics on educational outcomes, the results of these studies cannot be compared.

With respect to parental separation, however, it seems that differences between countries only play a small role in influencing the effects of parental separation on educational outcomes. A negative association between parental separation and educational outcomes has been observed in such different countries as the United States (McLanahan and Sandefur 1994; Sandefur and Wells 1999), England (Ermisch and Francesconi 2001b), and Sweden (Jonsson and Gähler 1997). In line with these results of previous research I find little difference in parental separation effects and their variation with social origin between Germany (Chapter 3) and Sweden (Chapter 4).

The analysis of German data is a peculiarity because Germany was separated into two states until reunification in 1990. Kesler (2003) shows that these legacies have implications for social origin inequalities in educational outcomes. She finds that, for recent cohorts, such as the German data I use, educational inequalities by social origin are lower in East than in West Germany. There are also differences between East and West Germany in terms of demographic outcomes such as fertility (Buhr and Huinink 2015; Goldstein and Kreyenfeld 2011), in the intergenerational transmission of divorce (Engelhardt et al. 2002), and children's experience of parental separation (see Chapter 3).

Given these important differences between East and West Germany, I have analyzed all German results on the full sample as well as on a sample without those children whose parents lived in the GDR in 1989. In addition, I also conducted all analyses on a sample restricted to the children from families with a GDR background. However, due to the low number of respondents these last models led to very imprecise estimates, which is why I do not report them. I base my conclusions on differences in the effects of family and sibling characteristics on educational outcomes on the comparison of the full sample and the sample which excludes children from families with a GDR background.

Given the objective of this thesis, it is not my aim to explain any cross-country differences in effects. I also do not see reasons to expect that the effects of the life events and characteristics under study, and their variation with social origin, vary between countries. Furthermore, the focus of the analysis, as will be detailed in the next section, is on measures of educational performance. Social origin inequalities in educational performance have been found to vary only little between countries (Jackson and Jonsson 2013). For these reasons, I do not expect the compensatory effect of social origin to vary between countries, although,

obviously, I cannot rule out that there are cross-country differences. The possible variation of the compensatory effect between different countries is an issue further research may focus on.

### **Children's Educational Outcomes**

This thesis analyzes the effects of month of birth, parental separation, birth order, birth spacing, and maternal age in England, Germany, and Sweden. The chapters differ in the educational outcomes they analyze. The outcomes under study in each chapter are chosen based on theoretical reasons and taking into account the particular characteristics of the educational systems under study. My main aim was to analyze in each country the most important educational outcomes for the further life chances of the children living in these countries.

Research has shown that socio-economic inequalities in cognitive abilities develop at an early age (Feinstein 2003). These early differences influence outcomes later in life, such as income (Almond and Currie 2011; Cunha and Heckman 2007; Heckman 2007). Furthermore, in most educational systems, performance is measured at different ages of the children's school careers. These measures, school grades or test scores, are certainly influenced by family and sibling characteristics as well as by social origin and are therefore important measures of children's educational outcomes. Test scores and school grades do differ in the way that school grades may be influenced by the subjective evaluation of performance of the teachers. This evaluation may differ between social classes. There is, however, no reason why the socio-economic heterogeneity in the effects of the life events and characteristics under study in this thesis should differ because of teacher influences on grading.

In most countries transition to the next or to a higher track in the education system depends on children's performance in school. In addition to their impact on educational performance, family and sibling characteristics and social origin may, however, also impact the range of educational options considered viable. The idea that social origin is associated with educational attainment even net of abilities has become popular in European educational sociology under the label "secondary effects" (Boudon 1974; Erikson et al. 2005; Jackson et al. 2007; Jackson 2013). These choices may depend on the education system; and, therefore, social origin inequalities in educational decision making may show greater cross-country variation than social origin inequalities in educational performance (Jackson and Jonsson 2013).

In this thesis, I mostly employ measures of educational performance because I expect the life events and characteristics I study to influence primarily educational performance. If they exercise effects on educational choices, this may largely be mediated through their influences on educational performance. However, because of the endogeneity of educational performance and educational choices to each other, the mediation hypothesis cannot be directly tested.

Instead, I assume that, when I use educational performance and educational attainment as separate outcome variables, mediation takes place if I find similar results of the effects of a characteristic on both types of educational outcomes. Consequently, I argue that the family and sibling characteristics under study have a different causal effect on educational performance and educational choices if the influence of these characteristics is different for measures of educational performance than for measures of educational transitions.

In the analysis on England and Germany I also employ measures of tracking. In the English case, I employ a measure of the decision to continue with an academic track after age 16. In Germany, I measure whether the respondents attend the highest track in the German education system when they are 16 to 17 years old (that is several years after which tracking took actually place).

Track attendance is a measure which is largely influenced by previous educational performance. However, it may also reflect choices in the range of educational options considered viable net of educational performance (Jackson 2013). Using an educational reform in one German state, Dollmann (2011) provides evidence that a system in which families can choose the track their children attend leads to higher educational inequalities than a system in which parents have to accept the recommendation of track attendance in secondary school made by the teachers of their children in primary school.

In general, I do not expect differences in the causal effects of life events under study on educational performance on school grades and on track attendance. But there is an exception; in the last empirical chapter in which I discuss that close birth spacing may have different effects on abilities and on choices. Contrary to that, the month of birth penalty is unlikely to influence educational decision making directly but more likely to have an impact on school grades which are, in the case of England, mostly based on centralized exams. Through its impact on school grades, month of birth does then also influence the decision to continue with academically oriented education at age 16 years. Parental separation may influence school grades if the child's school performance drops following this event. And it may also influence track attendance, but in the latter case the effect could also be brought about by a drop in

performance. Alternatively, educational decision making could be different in families with a non-resident parent. Since, however, I obtain similar results for the effect of parental separation on track attendance as well as on school grades, I rather assume the former to be the case.

In the first empirical study, our interest was to use as many measures of school grades as possible so that we could follow the month of birth penalty and the variation of this penalty by family background during the school career. Based on this consideration we actually chose both the country under study as well as the data sets which we employed. One feature of the English education system is that the performance of students is tested at several stages in the school career, the Key Stage results. The month of birth penalty is likely to mainly influence educational performance. For that reason we concentrate on measures of educational performance but we show that month of birth also affects the decision to continue with academic education after age 16 years.

The second empirical study uses information on school grades in Germany as well as a measure of track attendance. I argue in this chapter that track attendance is the most important outcome in the German education system (Hillmert and Jacob 2010). Nevertheless, I think it is useful to include further educational outcomes in the analysis, if only as a sort of robustness check.

The third empirical study uses data on Sweden and employs *meritvärde*, a measure of grade point average (GPA) in ninth grade. I chose this outcome because previous research has shown that the huge majority of inequality in educational opportunity in Sweden is captured by GPA in ninth grade (Erikson and Rudolphi 2010; Rudolphi 2013). This is the point at which the first important transition in the Swedish education system takes place.

The fourth empirical study differs from the previous ones because it is of a more explorative nature, in particular in the first part of the paper which is descriptive. For that reason, I also included a measure of cognitive performance in this analysis. Surely there are other German data sets that have more complex measures of cognitive ability than the SOEP. However, these do not have information on siblings and are therefore not suited for the analysis of within-family inequality in education in Germany. The measure of cognitive performance that I use has been used in previous research on the intergenerational transmission of cognitive skills (Anger 2012).

Measurement error may always influence the results. In order to reduce the influence of measurement error I use some of the most high-quality and widely used data. If there is random measurement error in the data this should bias the estimates downwards, hence,

making estimates rather conservative. Nevertheless, I cannot rule out that non-random measurement error is present in the data and in the analyses conducted on the basis of these data.

### **The Measurement of Social Origin**

Parents may provide many different resources to their children which influence their educational outcomes. The aim of the thesis is not to separate the different roles played by different resources; indeed I believe that such an enterprise would be difficult to undertake given the endogeneity of different types of resources to each other. Nor does this thesis aim to identify causal effects of parental class, education, income, or wealth on children's outcomes but to investigate whether the causal effects of certain life events and family characteristics differ between lower and higher social origin families.

Given this aim, I could operationalize social origin in several different ways. The only important requirement is to use a measure of social origin which I expect to capture most of what may provide socio-economic advantage in education. I mainly rely on using parental education as the measure to define social origin (Blossfeld et al. 2015; Peffer 2008). I chose to measure social origin via parental education for the following six reasons.

First of all, research has shown that parental education is more positively associated with child education than other measures of social origin (Buis 2013; Bukodi and Goldthorpe 2012). Hence, this research argues that parental education is a better predictor of child education than parental class or income.

Second, parental education is arguably the most stable measure of parental resources throughout childhood. This is important to ensure that the measure of social origin employed is exogenous to the life event under study. For instance, parental income may change following parental separation whilst parental education is not likely to change after parents separate.

Third, parental education is, for most people, causally preceding other measures of parental resources such as class, income, and wealth.

Fourth, parental education is the variable with the least measurement error, I assume, and the variable that is least likely to have missing values in all data sets I employ in this thesis.

Fifth, parental education is a measure of social origin which is used in different academic disciplines such as economics and sociology. Therefore I hope that, by using such a measure, my work can contribute across disciplines.

Sixth, most of the empirical research I quote with respect to having shown that the effects of certain life events and family characteristics vary by social origin, uses parental education to define social origin.

Although I use parental education to measure social origin throughout this thesis I do not make any causal claim about the effect of parental education. For instance, my finding is that the parental separation penalty differs by social origin. But I do not argue that this difference is brought about by parental education. In fact, it could be brought about by parental class, income, wealth, or any other parental characteristics associated with parental education. Such an analysis of which parental resource drives the compensatory effect is left to future research. It is not going to be an easy enterprise, however, since it requires at least two causal identification strategies in one study: first, to identify the causal effect of a family or sibling characteristic and, second, to identify the causal effect of the parental resource of interest, e.g. parental education. In that case the interaction effect between the family or sibling characteristic and the measure of social origin could be interpreted as the causal modification of the causal effect of the characteristic under study on child education.

Most chapters in this thesis distinguish two levels of parental education. This is done for both theoretical and practical reasons. A theoretical consideration is that in the countries I study in this thesis it is easy to split the sample into two parental education groups based on a crucial transition in the education system. Hence, a two-group definition of parental education can be quite easily applied in different countries allowing me to make the results comparable between countries. In Germany, I base the distinction between low and high social origin on whether at least one parent received *Abitur*, in England whether at least one parent obtained *A Level*, and in Sweden whether at least one parent attends the longer track in *Gymnasium*.

If there is strong heterogeneity this will be reflected in the estimates of the lower and the higher social origin group. For instance, if the difference is strongest between the top 10% and the lower 10% of the education distribution this will influence the estimates since the top 10% are included in the high social origin and the lower 10% are included in the low social origin group. Empirically, I show this in a robustness check which uses three categories of parental education in Chapter 2. This robustness check produces similar results as the main specification used in this and the other chapters which splits the sample into two parental education groups.

An advantage of the comparison between two groups is that just using two groups makes the estimates as precise as possible. Given the small part of the variance used in the causal identification strategies applied in this thesis this is a practical concern which cannot be neglected.

To sum up, using a two-level approach and parental education as a measure of social origin carries the advantage of distinguishing two groups of social origin based on a simple but powerful distinction which is also highly comparable across countries. Of course, there are many alternative ways to define social origin and in the robustness checks I explore some of them, for instance, by using parental class and parental ISEI as alternative measures of social origin in Chapter 5.

## **Overview of the Four Empirical Studies**

The four empirical studies collected in this thesis share the theoretical framework of the compensatory effect of social origin hypothesis. In addition, they share the methodological approach of applying research designs which control for unobserved heterogeneity. In total, the effects of five different variables and their variation by family socio-economic background are analyzed. I end this introduction by giving a brief overview of each of these four studies.

### *CHAPTER 2: MAKING UP FOR AN UNLUCKY MONTH OF BIRTH IN SCHOOL: CAUSAL EVIDENCE ON THE COMPENSATORY ADVANTAGE OF FAMILY BACKGROUND IN ENGLAND*

Previous research has shown that being born in the months immediately preceding the school entry cut-off date leads to lower educational outcomes in countries with a strict admission policy. In this chapter, we use the effect of age at school entry in England as an identification device to provide a causal estimate of the compensatory advantage enjoyed by children from high social origin families. We find that the negative effects of a young school entry age are stronger for children from low social origin families. We also investigate when social origin differences in school entry age effects emerge, and test possible mechanisms. We find that before starting school, a younger school entry age leads to lower test scores for children of both low and highly educated families. For children from highly educated families, the negative effect, however, progressively reduces over the school career and almost vanishes by age 16 years. With respect to the mechanisms underlying this compensatory effect, we find no

strong mediating role for parental involvement in homework and private lessons or for school choice.

*CHAPTER 3: WHEN GROWING UP WITHOUT A PARENT DOES NOT HURT: PARENTAL SEPARATION AND THE COMPENSATORY EFFECT OF SOCIAL ORIGIN*

This chapter investigates how the negative impact of parental separation on children's educational outcomes varies with social origin. In particular, I test the compensatory effect of social origin hypothesis which postulates that higher social origin families compensate the negative effects of disadvantageous life events, such as parental separation. I apply family-fixed effects models to control for unmeasured confounding characteristics of families and use data on siblings from Germany. I do find indication of substantial negative effects of parental separation on the probability of attending the upper track in secondary school (*Gymnasium*) and on school grades in German and Mathematics. These negative consequences of parental separation are limited to children in families where the parents have a low level of education. Children in families with highly educated parents are not negatively affected by their parents' separation in their educational outcomes. This finding supports the compensatory effect of social origin hypothesis and demonstrates that research on the consequences of parental separation has to take into account the heterogeneity of separation effects.

*CHAPTER 4: THE CAUSAL EFFECT OF PARENTAL SEPARATION ON CHILD EDUCATION: A NEW INSTRUMENTAL VARIABLE APPROACH*

This chapter proposes a new method to estimate the causal effect of parental separation on children's educational outcomes: using the ratio of opposite sex co-workers at the maternal workplace as an instrumental variable (IV) for parental separation. We apply this IV approach to Swedish pupils in order to estimate the effect of parental separation on their grade point average (GPA) at the end of primary school. We find that parental separation has, on average, no negative effect on GPA at the end of primary school in Sweden. Analyzing heterogeneity of separation effects, we find no differences in separation effects for boys and girls. However, we observe a negative causal effect of parental separation on child education in families with a low level of parental education. Contrary to that, no negative effect of parental separation is found in families with highly educated parents. Together with the results of Chapter 3, this

finding gives support to the compensatory effect of social origin hypothesis with respect to parental separation.

#### *CHAPTER 5: COMPETITION IN THE FAMILY: ESTIMATING AND EXPLAINING EDUCATIONAL INEQUALITIES BETWEEN SIBLINGS*

In this chapter I analyze inequalities in educational outcomes between siblings in Germany. I provide estimates for the proportion of educational inequalities which is produced within as opposed to between families. In particular, I test whether the relation between between- and within-family inequalities varies by family socio-economic background. Furthermore, I test which characteristics influence inequality between siblings, using family-fixed effects models. These characteristics include birth order, birth spacing, and maternal age. I find that differences in educational outcomes between siblings mostly exist in low- and high-SES families to a similar degree. In both, about half of educational inequalities are due to differences between siblings. Only for school grades I find a higher sibling similarity in high- than in low-SES families. In addition, I test whether the effects of sibling characteristics vary by family socio-economic background. In line with the compensatory effect of social origin hypothesis, the negative effect of having more closely spaced siblings is concentrated in low-SES families. On the other hand, the birth order and maternal age effects do not vary by family socio-economic background in a systematic way. These results give partial support to the notion that educational inequalities are misunderstood by not taking into account how different processes within families are in low- and high-SES families.

#### REFERENCES

- Akerlof, G. (1997). Social distance and social decisions. *Econometrica*, 65, 1005–1027.
- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124, 1729–1772.
- Almond, D. and Currie, J. (2011). Human capital development before age five. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics, Volume 4, Part B* (pp. 1315–1486). Amsterdam: Elsevier.
- Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annual Review of Economics*, 5, 37–56.

- Anger, S. (2012). Intergenerational transmission of cognitive and noncognitive skills. In J. Ermisch, M. Jäntti, and T. Smeeding (Eds.), *From Parents to Children: The Intergenerational Transmission of Advantage* (pp. 393–421). New York, NY: Russell Sage Foundation.
- Amato, P. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family*, 72, 650–666.
- Andrew, M. (2014). The scarring effects of primary-grade retention? A study of cumulative advantage in the educational career. *Social Forces*, 93, 653–685.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24, 3–30.
- Angrist, J. D. and Pischke, J.-S. (2015). *Mastering Metrics: The Path from Cause to Effect*. Princeton, NJ: Princeton University Press.
- Asbury, K. and Plomin, R. (2014). *G is for Genes: The Impact of Genetics on Education and Achievement*. Chichester, West Sussex: John Wiley & Sons.
- Ballarino, G., Bernardi, F., Requena, M., and Schadee, H. (2009). Persistent inequalities? Expansion of education and class inequality in Italy and Spain. *European Sociological Review*, 25, 123–138.
- Becker, R. (2003). Educational expansion and persistent inequalities of education: Utilizing subjective expected utility theory to explain increasing participation rates in upper secondary school in the Federal Republic of Germany. *European Sociological Review*, 19, 1–24.
- Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, 121, 1437–1472.
- Bell, D. (1973). *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. New York, NY: Basic Books.
- Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: A regression discontinuity based on month of birth. *Sociology of Education*, 87, 74–88.
- Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more, the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, 120, 669–770.

- Blake, J. (1985). Number of siblings and educational mobility. *American Sociological Review*, 50, 84–94.
- Blake, J. (1989). *Family Size and Achievement*. Berkeley, CA: University of California Press.
- Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. New York, NY: Wiley.
- Blossfeld, P. N., Blossfeld, G. J., and Blossfeld, H.-P. (2015). Educational expansion and inequalities in educational opportunity: Long-term changes for East and West Germany. *European Sociological Review*, 31, 144–160.
- Boudon, R. (1974). *Education, Opportunity, and Social Inequality*. New York, NY: Wiley & Sons.
- Bourdieu, P. (1984). *Distinction: A Social Critique of the Judgment of Taste*. Cambridge, MA: Harvard University Press.
- Bourdieu, P. and Passeron, J.-C. (1977). *Reproduction in Education, Society and Culture*. London: Routledge.
- Bozick, R., Alexander, K., Entwisle, D., Dauber, S., and Kerr, K. (2010). Framing the future: Revisiting the place of educational expectations in status attainment. *Social Forces*, 88, 2027–2052.
- Breen, R. and Goldthorpe, J. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31, 223–243.
- Breen, R., Luijkx, R., Müller, W., and Pollak, R. (2009). Non-persistent inequality in educational attainment: Evidence from eight European countries. *American Journal of Sociology*, 114, 1475–1521.
- Breen, R., Luijkx, R., Müller, W., and Pollak, R. (2010). Long-term trends in educational inequality in Europe: Class inequalities and gender differences. *European Sociological Review*, 26, 31–48.
- Breen, R., van de Werfhorst, H. G., and Jæger, M. M. (2014). Deciding under doubt: A theory of risk aversion, time discounting preferences, and educational decision-making. *European Sociological Review*, 30, 258–270.
- Buckles, K. S. and Munnich, E. L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources*, 47, 613–642.

- Buis, M. L. (2013). The composition of family background: The influence of the economic and cultural resources of both parents on the offspring's educational attainment in the Netherlands between 1939 and 1991. *European Sociological Review*, 29, 593–602.
- Buhr, P. and Huinink, J. (2015). The German low fertility: How we got there and what we can expect for the future. *European Sociological Review*, 31, 197–210.
- Bukodi, E. and Goldthorpe, J. H. (2012). Decomposing 'social origins': The effects of parents' class, status, and education on the educational attainment of their children. *European Sociological Review*, 28, 1–16.
- Burdick-Will, J. (2013). School violent crime and academic achievement in Chicago. *Sociology of Education*, 86, 343–361.
- Caudillo, M. L. and Torche, F. (2014). Exposure to local homicides and early educational achievement in Mexico. *Sociology of Education*, 87, 89–105.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartl, J., Mood, A. M., Weinfeld, F. D., and York, R.L. (1966). *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.
- Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York, NY: Pantheon Books.
- Conley, D. (2008). Bringing sibling differences in: Enlarging our understanding of transmission of advantage in families. In D. Conley and A. Lareau (Eds.), *Social Class: How Does It Work?* (pp. 179–200). New York, NY: Russell Sage Foundation.
- Conley, D. and Bennett, N. G. (2001). Birth weight and income: Interactions across generations. *Journal of Health and Social Behavior*, 42, 450–465.
- Conley, D., Domingue, B. W., Cesarini, D., Dawes, C., Rietveld, C. A., and Boardman, J. D. (2015). Is the effect of parental education on offspring biased or moderated by genotype? *Sociological Science*, 2, 82–105.
- Conley, D. and Glauber, R. (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, 722–737.
- Cunha, F. and Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97, 31–47.
- Davies, R., Heinesen, E., and Holm, A. (2002). The relative risk aversion hypothesis of educational choice. *Journal of Population Economics*, 15, 683–713.

- De Graaf, N. D., de Graaf, P., and Kraaykamp, G. (2000). Parental cultural capital and educational attainment in the Netherlands: A refinement of the cultural capital perspective. *Sociology of Education*, 73, 92–111.
- DiMaggio, P. (1982). Cultural capital and school success: The impact of status culture participation on the grade of U. S. high school students. *American Sociological Review*, 47, 189–201.
- DiPrete, T. A. and Eirich, G. M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, 32, 271–297.
- DiPrete, T. A. and Hout, M. (2006). What we have learned: RC28's contributions to knowledge about social stratification. *Research in Social Stratification and Mobility*, 24, 1–20.
- Dollmann, J. (2011). Verbindliche und unverbindliche Grundschulempfehlungen und soziale Ungleichheiten am ersten Bildungsübergang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 63, 595–621.
- Downey, D. (1995). When bigger is not better: Family size, parental resources, and children's wellbeing. *American Sociological Review*, 60, 746–761.
- Engelhardt, H., Trappe, H., and Dronkers, J. (2002). Differences in family policy and the intergenerational transmission of divorce: A comparison between the former East and West Germany. *Demographic Research*, 6, 295–324.
- Erikson, R., Goldthorpe, J., Jackson, M., Yaish, M., and Cox, D. R. (2005). On class differentials in educational attainment. *Proceedings of the National Academy of Sciences*, 102, 9730–9733.
- Erikson, R. and Jonsson, J. O. (1996). Explaining class inequality in education: The Swedish test case. In R. Erikson and J. O. Jonsson (Eds.), *Can Education Be Equalized? The Swedish Case in Comparative Perspective* (pp. 1–64). Boulder, CO: Westview Press.
- Erikson, R. and Rudolphi, F. (2010). Change in social selection to upper secondary school: Primary and secondary effects in Sweden. *European Sociological Review*, 26, 291–305.
- Ermisch, J. and Francesconi, M. (2001a). Family matters: Impacts of family background on educational attainments. *Economica*, 68, 137–156.
- Ermisch, J. and Francesconi, M. (2001b). Family structure and children's achievement. *Journal of Population Economics*, 14, 249–270.
- Ermisch, J. and Francesconi, M. (2013). The effect of parental employment on child schooling. *Journal of Applied Econometrics*, 28, 796–822.

- Esser, H. (1996). What is wrong with 'variable sociology'? *European Sociological Review*, 12, 159–166.
- Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica*, 70, 73–98.
- Freedman, D. (2004). Graphical models for causation, and the identification problem. *Evaluation Review*, 28, 267–293.
- Gabay-Egozi, L., Shavit, Y., and Yaish, M. (2010). Curricular choice: A test of a rational choice model of education. *European Sociological Review*, 26, 447–463.
- Gambetta, D. (1987). *Were they Pushed or Did they Jump? Individual Decision Mechanisms in Education*. Cambridge: Cambridge University Press.
- Gangl, M. (2010). Causal inference in sociological research. *Annual Review of Sociology*, 36, 21–47.
- Gelman, A. and Imbens, G. (2013). Why ask why? Forward causal inference and reverse causal questions. *National Bureau of Economic Research Working Paper*, No. 19614.
- Giesselmann, M. and Windzio, M. (2014). Paneldaten in der Soziologie: Fixed Effects Paradigma und empirische Praxis in Panelregression und Ereignisanalyse. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 95–113.
- Goldberger, A. (1979). Heritability. *Economica*, 46, 327–347.
- Goldstein, J. R. and Kreyenfeld, M. R. (2011). Has East Germany overtaken West Germany? Recent trends in order-specific fertility. *Population and Development Review*, 37, 453–472.
- Goldthorpe, J. H. (2001). Causation, statistics, and sociology. *European Sociological Review*, 17, 1–20.
- Goldthorpe, J. H. (2007). *On Sociology*. Palo Alto, CA: Stanford University Press.
- Goldthorpe, J. H. (in press). *Sociology as a Population Science*. Cambridge: Cambridge University Press.
- Guo, G. and VanWey, L. (1999). Sibship size and intellectual development. *American Sociological Review*, 64, 207–231.
- Grigg, J. (2012). School enrollment changes and student achievement growth: a case study in educational disruption and continuity. *Sociology of Education*, 85, 388–404.
- Hanselman, P., Bruch, S. K., Gamoran, A., and Borman, G. D. (2014). Threat in context: school moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education*, 87, 106–124.

- Hanushek, E. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100, 84–117.
- Hasan, S. and Bagde, S. (2013). The mechanics of social capital and academic performance in an Indian college. *American Sociological Review*, 78, 1009–1032.
- Härkönen, J. (2014). Birth order effects on educational attainment and educational transitions in West Germany. *European Sociological Review*, 30, 166–179.
- Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104, 13250–13255.
- Hillmert, S. and Jacob, M. (2010). Selections and social selectivity on the academic track: A life course analysis of educational attainment in Germany. *Research in Social Stratification and Mobility*, 28, 59–76.
- Jackson, M. (Ed.) (2013). *Determined to Succeed? Performance versus Choice in Educational Attainment*. Stanford, CA: Stanford University Press.
- Jackson, M. and Jonsson, J. O. (2013). Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context. In Jackson M. (Ed.), *Determined to Succeed? Performance versus Choice in Educational Attainment* (pp. 306–338). Stanford, CA: Stanford University Press.
- Jackson, M., Erikson, R., Goldthorpe, J. H., and Yaish, M. (2007). Primary and secondary effects in class differentials in educational attainment. *Acta Sociologica*, 50, 11–29.
- Jæger, M. M. (2008). Do large sibships really lead to lower educational attainment? New evidence from quasi-experimental variation in couples' reproductive capacity. *Acta Sociologica*, 51, 217–235.
- Jæger, M. M. (2011). Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data. *Sociology of Education*, 84, 281–298.
- Jencks, C. (1980). Heredity, environment, and public policy reconsidered. *American Journal of Sociology*, 45, 723–736.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., and Michelson, S. (1972). *Inequality: A Reassessment of Family and Schooling in America*. New York: Basic Books.
- Jonsson, J. and Gähler, M. (1997). Family dissolution, family reconstitution and children's educational careers: Recent evidence for Sweden. *Demography*, 34, 277–293.
- Kalmijn, M. and Kraaykamp, G. (2005). Late or later? A sibling analysis of the effect of maternal age on children's schooling. *Social Science Research*, 34, 634–650.

- Karlson, K. B. (2011). Multiple paths in educational transitions: A multinomial transition model with unobserved heterogeneity. *Research in Social Stratification and Mobility*, 29, 323–341.
- Kerr, C., Dunlop, J., Harbison, F., and Myers, C. (1960). *Industrialism and Industrial Man: The Problem of Labor and Management in Economic Growth*. Cambridge, MA: Harvard University Press.
- Kesler, C. (2003). Educational stratification and social change: Evidence from German unification. *European Sociological Review*, 19, 467–482.
- Kraphol, E., Rimfeld, K., Shakeshaft, N., Trzaskowski, M., McMillan, A., Pingault, J.-B., Asbury, K., Harlaar, N., Kovas, Y., Dale, P. S., and Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Science*, 111, 15273–15278.
- Kurlaender, M. and Grodsky, E. (2013). Mismatch and the paternalistic justification for selective college admissions. *Sociology of Education*, 86, 294–310.
- Lareau, A. (2003). *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.
- Lareau, A. and Weininger, E. B. (2003). Cultural capital in educational research: A critical assessment. *Theory and Society*, 32, 567–606.
- Lauen, D. L. and Gaddis, S. M. (2013). Exposure to classroom poverty and test score achievement: Contextual effect or selection? *American Journal of Sociology*, 118, 943–979.
- Legewie, J. and DiPrete, T. (2012). School context and the gender gap in educational achievement. *American Sociological Review*, 77, 463–485.
- Manski, C. (1995). *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.
- Manski, C. (2013). *Public Policy in an Uncertain World: Analysis and Decisions*. Cambridge, MA: Harvard University Press.
- Manzo, G. (2013). Educational choices and social interactions: a formal model and a computational test. *Comparative Social Research*, 30, 47–100.
- Mare, R. D. and Tzeng, M.-S. (1989). Fathers' ages and the social stratification of sons. *American Journal of Sociology*, 95, 108–131.
- Marteletto, L. J. and de Souza, L. R. (2013). The implications of family size for adolescents' education and work in Brazil: Gender and birth order differences. *Social Forces*, 92, 275–302.

- McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography*, 41, 607–627.
- McLanahan, S. and Sandefur, G. (1994). *Growing Up with a Single Parent: What Hurts, What Helps?* Cambridge, MA: Harvard University Press.
- McLanahan, S., Tach, L., and Schneider, D. (2013). The causal effects of father absence. *Annual Review of Sociology*, 39, 399–427.
- Moffitt, R. A. (2005). Remarks on the analysis of causal relationships in population research. *Demography*, 42, 91–108.
- Morgan, S. L. (2005). *On the Edge of Commitment: Educational Attainment and Race in the United States*. Stanford, CA: Stanford University Press.
- Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research. Second Edition*. Cambridge: Cambridge University Press.
- Pettersson-Lidbom, P. and Skogman Thoursie, P. (2009). Does child spacing affect children's outcomes? Evidence from a Swedish reform. *Institute for Labour Market Policy Evaluation (IFAU) Working Paper*, 2009-7.
- Pfeffer, F. (2008). Persistent inequality in educational attainment and its institutional context. *European Sociological Review*, 24, 543–565.
- Plomin, R., DeFries, J. C., Knopik, V. S., and Neiderhiser, J. M. (2013). *Behavioral Genetics. Sixth Edition*. New York, NY: Worth Publishers.
- Powell, B. and Steelman, L. C. (1993). The educational benefits of being spaced out: Sibship density and educational progress. *American Sociological Review*, 58, 367–381.
- Price, J. (2008). Parent-child quality time: does birth order matter? *Journal of Human Resources*, 43, 240–265.
- Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J. et al. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, 340, 1467–1471.
- Rudolphi, F. (2013). Ever declining inequalities? Transitions to upper secondary and tertiary education in Sweden, for the 1972-1990 birth cohorts. In M. Jackson (Ed.), *Determined to Succeed* (pp. 185–227). Palo Alto, CA: Stanford University Press.
- Sandefur, G. and Wells, T. (1999). Does family structure really influence educational attainment? *Social Science Research*, 28, 331–357.
- Sewell, W. H., Haller, A. O., and Portes, A. (1969). The educational and early occupational attainment process. *American Sociological Review*, 34, 82–92.

- Sewell, W. H. and Hauser, R. M. (1975). *Education, Occupation, and Earnings: Achievement in the Early Career*. New York, NY: Academic Press.
- Sewell, W. H. and Hauser, R. M. (1980). The Wisconsin Longitudinal Study of social and psychological factors in aspirations and achievements. *Research in Sociology of Education and Socialization*, 1, 59–99.
- Shavit, Y. and Blossfeld, H.-P. (1993). *Persistent Inequality: A Comparative Study of Educational Attainment in Thirteen Countries*. Boulder, CO: Westview.
- Steelman, L., Powell, B., Werum, R., and Carter, S. (2002). Reconsidering the effects of siblings configurations: Recent advances and challenges. *Annual Review of Sociology*, 28, 243–269.
- Stocké, V. (2007). Explaining educational decision and effects of families' social class position: an empirical test of the Breen-Goldthorpe model of educational attainment. *European Sociological Review*, 23, 505–520.
- Tieben, N., de Graaf, P. M., and de Graaf, N. D. (2010). Changing effects of family background on transitions to secondary education in the Netherlands: consequences of educational expansion and reform. *Research in Social Stratification and Mobility*, 27, 77–90.
- Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40, 1008–1018.
- Treiman, D. (1970). Industrialization and social stratification. In E. Laumann (Ed.), *Social Stratification: Research and Theory for the 1970s* (pp. 207–234). Indianapolis, IN: Bobbs-Merrill.
- Watts, D. (2015). Common sense and sociological explanations. *American Journal of Sociology*, 120, 313–351.
- Willis, P. (1977). *Learning to Labour*. Farnborough: Saxon House.
- Winhsip, C. and Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25, 659–707.
- Zajonc, R. B. and Markus, G. B. (1975). Birth order and intellectual development. *Psychology Review*, 37, 1325–1341.

## **CHAPTER 2: MAKING UP FOR AN UNLUCKY MONTH OF BIRTH IN SCHOOL: CAUSAL EVIDENCE ON THE COMPENSATORY ADVANTAGE OF FAMILY BACKGROUND IN ENGLAND<sup>1</sup>**

### **Introduction**

In nearly all education systems, children enter school the year in which they reach a certain age before a given cut-off date. Since children are born in different months during the year, children in the same class differ substantially in age at school entry. The youngest children in a class are almost twelve months younger than the oldest and previous research has shown that children who are younger upon entering school tend to have substantially worse educational outcomes (Bedard and Dhuey 2006; Black et al. 2011; Crawford et al. 2010; Crawford et al. 2013; Dobkin and Ferreira 2010; Fredriksson and Öckert 2014; Mühlenweg and Puhani 2010).

Since postponement of school entry is not allowed in England, under the testable assumption that month of birth is not associated with family background, we use age at school entry as a randomly allocated characteristic and investigate whether its negative effect on education is smaller for children from high social origin families. This result is predicted by the compensatory effect of social origin hypothesis according to which a disadvantageous life event has less negative consequences in high than in low social origin families (Bernardi 2014).

In addition, we investigate when the compensatory advantage enjoyed by children from a socioeconomically privileged background occurs. We are interested in finding out whether the compensation for a disadvantageous month of birth has already occurred before school start or whether it occurs later in the school career. Finally, we test three possible mechanisms underlying the compensatory advantage: private lessons, parental help with homework, and school choice.

We use data from two cohort studies in England: the Millennium Cohort Study (MCS) and the Longitudinal Study of Young People in England (LSYPE). These data allow us to

---

<sup>1</sup> This chapter was co-authored with Fabrizio Bernardi. A shortened version of this chapter was published in 2015 as “Making Up for an Unlucky Month of Birth in School: Causal Evidence on the Compensatory Advantage of Family Background in England.” *Sociological Science*, 2, 235-251. DOI: 10.15195/v2.a12. We are grateful to The Centre for Longitudinal Studies, UCL Institute of Education and the Department for Education/ National Centre for Social Research for the use of the data in this chapter and to the UK Data Archive and Economic and Social Data Service for making them available. However, they bear no responsibility for the analysis or interpretation of these data.

measure educational performance at various points in a school career and to follow the evolution of social origin differences in school entry age effects over time.

## **Background**

### Month of Birth and School Entry Age Effect

Previous studies have discussed three explanations for the observed negative effect of a young school entry age on educational outcomes (Bedard and Dhuey 2006). First of all, through the cut-off date children differ in their school entry age as well as in the age at which they sit exams. Children with a later school entry age might perform better on tests because they are older than their peers when exams are taken. A student born in September is, on average, eleven months older than a student born in August when they sit a test on the same date. The older student has, therefore, had more time to learn and to develop his or her cognitive abilities. This explanation is usually referred to as a “maturity effect” or as an “absolute age effect” (Crawford et al. 2014).

The second mechanism points to a peer effect. Older children perform better in class because they are physically, emotionally, and intellectually more developed than their peers when they start school. This initial relative advantage due to their month of birth affects their self-esteem and motivation, which can have enduring effects on later performance.

The third mechanism points to an age-specific learning effect. The life cycle skill formation model suggests that an early advantage in skills facilitates acquiring further skills (Cunha et al. 2006; Cunha and Heckman 2007). Older children upon entering school are more ready to learn and to profit from school. This initial advantage puts them on a positive trajectory with long-term consequences for educational achievement.

In England there is a fourth reason why month of birth may have an impact on educational outcomes. Due to the admission rules for the reception class (the first year of primary school in England), the amount of schooling children have received depends on their month of birth (Crawford et al. 2010). In some areas children born later in the academic year may be granted access to school one or two terms later than their older peers. This rule suggests an additional mechanism through which children born just before the cut-off date may be further penalized: they spend less time in school.

Both peer effect and life cycle skill formation model mechanism suggest that the negative effect of a young school entry age should persist and possibly become even stronger

over the course of an educational career. If a maturity effect is in place, however, we should expect that the effect of month of birth diminishes over the school career since the relative weight of age differences upon entering school becomes smaller as children get older. The eleven months difference between a student born in August and one born in September becomes progressively less significant at older ages. For instance at age six years, being eleven months older provides a 15 percent advantage in terms of lived months (11/72), while at age 16 years the advantage is reduced to five percent (11/192). Similarly, the advantage associated with earlier admission to the reception class should also become smaller as children grow older.

Identifying the role played by each of these mechanisms is, however, complicated as they may interact. The empirical evidence from different countries is inconclusive (Black et al. 2011; Crawford et al. 2013; Dobkin and Ferreira 2010; Fredriksson and Öckert 2013; Mühlenweg and Puhani 2010). The most recent empirical evidence for England suggests that the main driver of the month of birth penalty is the maturity effect (Crawford et al. 2014).

#### Heterogeneity of Month of Birth Effects and the Compensatory Advantage of Social Origin

Compensatory advantage describes the notion that children from socioeconomically advantaged families are more sheltered from the long-term consequences of prior disadvantageous life events and characteristics that negatively influence educational outcomes (Bernardi 2014; Bernardi and Cebolla-Boado 2014). A disadvantageous characteristics or life event may impact negatively the school career for children coming from low social origin families while the same disadvantage may be less detrimental for the educational outcomes of children coming from high social origin families.

Two basic tenets in social stratification research suggest that a compensatory advantage is likely to be observed whenever a disadvantageous life event or characteristic endangers the chances of future educational and occupational achievement. First, families aim for their offspring to achieve at least their own social position (Boudon 1998; Breen and Goldthorpe 1997). In this respect, high social origin families have a stronger incentive to compensate for events and characteristics that endanger their children's future educational and occupational opportunities. Second, high social origin families have the financial, cultural, and social resources to pursue compensatory strategies. Based on the notion of compensatory advantage, our main hypothesis is that the negative implications of a young school entry age are concentrated among students from low social origin families.

So far causal evidence for such a compensatory advantage mainly comes from research on the effects of early health conditions on later educational and occupational outcomes. Almond and Mazumder (2013) provide a recent overview of this literature. Using exposure to radiation after the Chernobyl accident as a natural experiment, Almond et al. (2009) show that the effects of early health shocks on later outcomes only persist in families with low educated fathers. Similarly, Torche and Echevarría (2011) show that the negative consequences of a low birth weight are largely concentrated in families with a low level of maternal education. In two similar but independent studies, Hsin (2012) and Restrepo (2012) study parental investment responses to birth weight differences between siblings growing up in the same families as a possible mechanism to explain these findings. Both studies come to the conclusion that families with a high level of parental education invest more in the disadvantaged sibling with a lower birth weight whilst families with a low level of parental education invest more in that sibling who has a higher birth weight.

A second issue is the timing of when such a difference between social origin groups occurs. Studies of cognitive development in early childhood have consistently shown inequalities in cognitive skills before children enter school (Becker 2011), in pre-school attendance (Sylva et al. 2010), and performance differences within the first school year (Tymms et al. 1997). Following this line of research, one may expect that the compensation of the month of birth penalty has already taken place before school entry. According to this line of thought, high social origin families are aware of and anticipate the potential disadvantage associated with an early school entry age. They make sure that their children are ready to start school despite their young age. However, an alternative explanation is that children from a high social origin who are born just before the cut-off date for admission to primary education may catch up later in the school career. This would be the case if, instead of compensating before school start, parents actively reacted to the low performance of their children once it became manifest in school, and took some course of action to address it.

Finally, in our study we are able to test two sets of explanations for how the compensatory advantage may be brought about. To start with, there may be parental actions directly oriented towards improving school performance. Examples of such actions are parents helping with homework or paying for additional private lessons (Coleman 1988). In addition, earlier research has shown that the sorting of children into different schools has a strong impact on educational inequalities (Ermisch and Del Bono 2012). Compensation of the disadvantage associated with a young age at school start might, then, occur as a result of social origin differences in school choice if high social origin families enroll their children in

schools that are more effective in raising the performance of initially low-performing students.

## **The English Education System**

Education is compulsory in England until age 16. Thereafter students and their families must decide whether to continue to take Advanced Level education (*A Levels*), which allows students to later pursue a university degree; to pursue a vocational degree; or to leave the education system. Before this decisive point in the English education system, performance is measured in centralized exams at certain ages. These exams are taken at the end of so-called Key Stages. In the analysis we employ scores at Key Stage 2 (age 11), Key Stage 3 (age 14), and Key Stage 4 (age 16). The exams at the end of Key Stage 4 are the most important ones, since they lead to the General Certificate of Secondary Education (GCSE) or equivalent qualifications. Grades range from A\* to G, with A\* being the highest and G the lowest. The achievement of five grades in the A\* to C range is important for further educational attainment.

The cut-off date for admission to the first year of primary education (reception class) in England is the first of September. In most areas in England, children are admitted to reception class in September of the academic year in which they turn five (Crawford et al. 2010; Eurydice 2011). However, in other areas, pupils are admitted to reception class in the term in which they turn five. Children born between September and December are admitted in September, those born between January and April in January, and those born between May and August in April (Crawford et al. 2010). Under all circumstances, those born in August are the youngest in a school cohort, whether they start reception class in September or in a later term (January or April). In the latter case, they are even further disadvantaged by their month of birth because they receive less schooling. In England, contrary to what happens in other countries, postponement of admission to the first year of primary education is not allowed by educational authorities (Eurydice 2011).

## **Research Design**

In a recent article, Bernardi (2014) uses a regression discontinuity based on the school entry cut-off date in France and shows that French students born just before the cut-off date have a higher risk of grade retention than children born after the cut-off date. The risk is, however,

much smaller for such students who are born to highly educated parents than for those from families with a low level of parental education. In this article, we employ a similar research design for the English case, although we consider different educational outcomes and mainly employ a linear definition of school entry age.

To develop our arguments more formally, one can consider a regression model that predicts the negative influence of a given disadvantageous life event or characteristic  $C_i$  (examples may include a low birth weight, a high birth order, the experience of parental separation, etc.) on an educational outcome:

$$E_i = C_i \beta + X_i \gamma + \varepsilon_i \quad (1)$$

with  $E_i$  being the educational outcome of interest and  $X_i$  being a vector of control variables. The error term is written as  $\varepsilon_i$ . The subscript  $i$  describes the individual.

If we are now interested in heterogeneity by social origin, we add an interaction with an indicator of social origin to the model:

$$E_i = C_i \beta + SO_i \delta + C_i \times SO_i \zeta + X_i \gamma + \varepsilon_i \quad (2)$$

The interaction between  $C_i$  and  $SO_i$  tells how the association between a disadvantageous characteristic and an educational outcome varies by family background.

The endogeneity of  $C_i$  makes a causal interpretation of the parameters of Equation (2) problematic. If some unobserved variables affect both the disadvantageous characteristic  $C_i$  and the educational outcome  $E_i$ , both the estimate of  $\beta$  and the estimate of the interaction effect  $\zeta$  are biased and inconsistent. In order to test the compensatory effect of social origin hypothesis causally, the disadvantageous characteristic  $C_i$  must be exogenous. We argue that starting school at a young age in England (as in other countries with a strict cut-off date for admission) is such an exogenous disadvantageous trait. The model that will be estimated can, then, be written as:

$$E_i = A_i \beta + SO_i \delta + A_i \times SO_i \zeta + X_i \gamma + \varepsilon_i \quad (3)$$

with  $A_i$  being a continuous variable of school entry age, depending on the month of birth relative to the school cut-off date. If  $A_i$  is truly exogenous, the interaction effect  $A_i \times SO_i$  can be causally interpreted. Alternatively,  $A_i$  can be constructed as a dummy variable that

distinguishes those born in August, just before the cut-off date, and those born in September, just after it. In this second specification, Equation (3) becomes a regression discontinuity (Dobkin and Ferreira 2010). Whilst we deal with the continuous school entry age variable in the main text, we additionally provide a robustness check using the regression discontinuity definition.

The causal interpretation of the school entry age effect depends crucially on the assumption that month of birth is randomly distributed across the population. We will therefore start our empirical analysis precisely by testing this assumption. Before doing that we describe in the next section the data, variables, and models that we use.

## **Data, Variables, and Models**

### Data

We use data from two English cohort studies: the Millennium Cohort Study (MCS) (University of London 2010) and the Longitudinal Study of Young People in England (LSYPE) (Department for Education and National Centre for Social Research 2011). These survey datasets are representative of two English cohorts of children, with the LSYPE using a sample of children born around ten years earlier than the children in the MCS. However, due to restrictions in the educational outcomes available in the datasets, we use the MCS to cover the earlier years and the LSYPE to cover the later years of the children's school careers.

The MCS samples children who were born in 2000 and 2001. We restrict the sample to children born between September 2000 and August 2001 in order to have one cohort of children who entered school in the same academic year. We use information from the first three waves of the MCS.

The LSYPE samples children who were aged 13 to 14 in 2004 and follows them with yearly updates in seven waves until 2010. The survey started with an initial sample size of 15,770 participating pupils in 2004. The survey samples children attending both maintained and independent schools. Again, we restrict the sample to children born within one academic year cohort; in the case of this survey, between September 1989 and August 1990. The LSYPE data includes information on school grades from the National Pupil Database, allowing us to use information on school grades for these pupils at ages 11, 14, and 16.

We restrict both samples to pupils who were in England at the time of the survey, and the LSYPE sample to children born in the UK to make it more similar to the MCS data which sampled who lived in the UK nine months after their birth.

## Variables

The central independent variable in our analysis is relative age at school start (abbreviated as relative age in the tables). The variable is coded so that being born in the month immediately before the cut-off date (August) is equal to 0, while being born in September is equal to -11. This means that younger children upon entering school (those born in August) are given a higher value for this variable. Its coefficient can, therefore, be interpreted as the penalty of a younger school entry age. In addition to this linear specification, as a robustness check we consider age at school entry as a categorical variable and compare those born in the three months just before and the three months just after the cut-off date. The results are fully in line with those obtained using the linear school entry age variable.

We define social origin by the highest level of education achieved by a child's father or mother. The variable is coded dichotomously with a high level of parental education meaning that one of the parents has obtained *A Level*, an equivalent, or a higher qualification. Based on these definitions, between 44 percent (MCS) and 47 percent (LSYPE) of the children sampled have parents with a high level of education (see Table 1).

We analyze the effect of relative age on different measures of cognitive performance and educational outcomes. The following outcomes are estimated as dependent variables in regression models (the data source used to estimate the specific outcome is in parentheses):

- Scores on the British Ability Scale (BAS) at age five (MCS)
- Key Stage 2 scores at age 11 (LSYPE)
- Key Stage 3 scores at age 14 (LSYPE)
- Achievement of at least five GCSE (and equivalents) at grades between A\* to C at age 16 (LSYPE)
- Decision to continue with academically oriented education after age 16 (LSYPE)

Performance at Key Stages 2 and 3, as well as performance on the GCSE, is measured for all pupils at about the same time, but performance in the MCS is measured by tests taken in the third wave of the survey when the children were about age five, and the tests were not taken by all children on the same day. The fieldwork conducted during the third wave of the MCS took place between September and May 2006. Children born in September were

interviewed at the beginning of the fieldwork period and those born in August toward its end. However, since the fieldwork period was shorter than the time frame in which the children were born apart, the children born in August were, on average, younger than children born in September when the survey test was administered. For each month a respondent was born later in the academic year, on average, he or she was about 21 days younger when the survey test was taken. Because of this fieldwork design, some respondents were younger at the time of the test solely because of their month of birth. We argue that for that reason, the fieldwork design operates like a cut-off date for admission to school. It should be noted, however, that we are likely to underestimate the month of birth penalty at age five. The age difference, due to month of birth, for national tests taken on the same day at age 11, 14, and 16 is larger than the age difference at the time of the survey at age five. If the month of birth penalty is smaller, it should be easier for children from highly educated families to catch up. The analysis at age five using the MCS therefore provides a conservative test of the hypothesis that the compensatory advantage enjoyed by children of highly educated families is already in place before starting school.

The MCS assessment at age five includes three standardized tests measured by the British Ability Scale (BAS). We report the results for the Vocabulary score in the main text and the results for the Picture Similarity and the Pattern Construction scores in Table A2.1 in the appendix. There are no differences in results between the three test scores. We standardize these outcome variables with a mean of 0 and a standard deviation of 1 so that effects of the independent variables can be interpreted in terms of standard deviations.

At Key Stage 2 (age 11) and Key Stage 3 (age 14) performance is measured in Mathematics, English, and Science via standardized tests. At both Key Stages we employ an overall performance score as an average of these test scores. Test results come close to being normally distributed. These outcome variables are also standardized with a mean of 0 and a standard deviation of 1.

At Key Stage 4 (age 16), performance refers to the GCSE (and equivalent) exams. As a summary measure of performance on the GCSE exams we consider whether someone achieved a grade between A\* and C in at least five GCSE (and equivalent qualifications). In the appendix, we also report results of an OLS model predicting the total numbers of A\* to C in the GCSE exams. Finally, we investigate whether or not the respondents decide to continue with academic education after age 16. About 60 percent of students achieve five GCSE (and equivalents) with A\* to C grades or more, and about 73 percent continue with academic education (Table 2.1).

Table 2.1 reports descriptive statistics on both analysis samples.

**Table 2.1** Descriptive statistics

	Millennium Cohort Study (MCS)			Longitudinal Study of Young People in England (LSYPE)		
	Mean	Std. dev.	<i>N</i>	Mean	Std. dev.	<i>N</i>
Relative age	-5.53	3.47	9,277	-5.38	3.47	13,953
Male	0.51	0.50	9,277	0.51	0.50	13,823
High parental education	0.44	0.50	9,252	0.47	0.50	13,918
British Ability Scale, vocabulary score at age 5 <sup>1</sup>	106.30	17.03	9,117			
Key Stage 2 score <sup>1</sup>				26.99	4.02	13,221
Key Stage 3 score <sup>1</sup>				33.83	6.59	13,178
5 GCSE at grades A* to C				0.60	0.49	13,625
Academic education after 16				0.73	0.45	9,519
Private lessons				0.23	0.42	13,921
Help with homework				0.86	0.34	13,579

*Source:* Millennium Cohort Study (MCS), Longitudinal Study of Young People in England (LSYPE).

<sup>1</sup> We report means and standard deviations on these variables before standardization. In the models we, however, regress on these variables after standardization.

With regard to the mechanisms that possibly underlie the observed compensatory advantage, we use information provided by the parents on whether they paid for private lessons. At each wave of the LSYPE, parents report whether they have paid for their children to have extra lessons in the past twelve months. We use the first four waves of the LSYPE, which cover the phase between ages 13 and 16, to construct a dummy variable valued at 1 if parents report that they have paid for their son or daughter to have any extra lessons within the time period covered by the four survey waves. A similar dummy variable is constructed with respect to homework, with pupils reporting whether they had any support at home in completing their homework. This information was included in the first two waves of the LSYPE, which covers the period between ages 13 and 14. For that reason, help with homework is a dummy variable coded 1 if students report that they received support at home for completing their homework at age 13 or 14.

The role played by school choice in providing compensation to students with a young school entry age coming and highly educated parents is tested through comparison of the month of birth penalty within and between schools. We apply school-fixed effects models and compare the results of these models to the results of the cross-sectional estimates. School-fixed effects models use only the variation within schools by comparing pupils who attend the same school (Ermisch and Del Bono 2012). We argue that a reduction in the school entry age

effect in the school-fixed effects models would mean that the selection of school, which is influenced by family background, plays a role in explaining the compensatory advantage of students from highly educated families.

## Models

We estimate OLS regression models for test scores and Linear Probability Models (LPM) for the achievement of five GCSE with grades A\* to C and for the decision to continue with academically oriented education after age 16. We use LPM in the latter case because of the straightforward interpretation of their estimates, in particular the interaction effects which we focus on in our analysis (Angrist and Pischke 2009; Mood 2010). For all our estimates we present ten percent significance levels based on one-tailed tests because we have a clear hypothesis for both the direction of the school entry age effect and the compensatory effect (Freedman et al. 2007). Commenting on the findings, we concentrate on effect sizes.

## Results

### The Association between Month of Birth and Social Origin

We start by analyzing the association between month of birth and parental education in England. Buckles and Hungerman (2013) found that families from different social origins in the United States tend to give birth to children in different seasons of the year. The crucial assumption of our research design is, however, that there is no systematic sorting of birth dates by social origin in the months before and after the school entry cut-off date (i.e. before and after September 1). Previous studies conducted in England suggest that this assumption is valid (Crawford et al. 2010; Crawford et al. 2014). Nor do we find any sizeable differences between low and highly educated parents in the propensity to give birth in a given month.

We report results on the association between parental education and month of birth in Table 2.2. The chi-square tests for the association between parental education and month of birth are statistically insignificant and the Cramér's  $V$  measures of association are almost equal to zero in both the MCS and the LSYPE samples. It should be noted, that our samples are large enough to detect any substantively large effects.<sup>1</sup> Based on these results we conclude that there is no association between month of birth and parental education. This finding supports our strategy of interpreting month of birth as an exogenous explanatory variable.

## Variation in the Effects of School Entry Age on Educational Outcomes

Table 2.3 reports regression models of the effects of age at school entry on educational outcomes and their variation by social origin. The first model for each educational outcome documents the effect of school entry age at different stages of the school career. In line with previous research we find that being born in August instead of September entails a sizeable penalty in terms of educational achievement. For instance at age five, being one month younger leads to a 0.035 standard deviations lower Vocabulary score on the British Ability Scale. The result means that being born in August instead of September reduces the Vocabulary score by 0.385 standard deviations.

The month of birth penalty decreases between Key Stage 2 and Key Stage 3 but is still sizeable at the GCSE exams at age 16. Model 7 shows that the likelihood of receiving at least five GCSE (or equivalents) with grades between A\* and C is about seven percentage points ( $11 \times -0.006 = -6.6$  percent) lower for children born in August than for children born in September. The disadvantage based on month of birth is thus not trivial, and similar in size to gender inequality in educational outcomes, which has received substantial attention in recent research (Buchmann and DiPrete 2006). What is more, there is also a month of birth penalty for the decision to pursue further academically oriented qualifications, as Model 9 demonstrates.

The second set of models investigates whether the effect of relative age varies by parental education at different ages. Model 2 shows that having highly educated parents does not mitigate in any sizable way the disadvantage of an unlucky month of birth before starting school. In contrast, at ages 11 and 14 there are hints that a compensatory advantage starts to manifest itself. In fact, the interaction between parental education and relative age is positive and becomes more sizeable over time. It is, however, at age 16 years that the compensatory advantage becomes fully evident. For someone whose parents are not highly educated, being born in August implies a penalty of eleven percentage points ( $11 \times -0.010 = -11$  percent) in the probability of achieving at least five GCSE with grades A\* to C, when compared to someone born in September. For those whose parents are highly educated, the same penalty is only three percentage points ( $11 \times -0.010 + 11 \times 0.007 = -3.3$  percent). In the robustness check, we report results using two alternative outcome measures at age 16 years, which are continuous and can, therefore, be standardized. These results also suggest that the interaction between relative age and parental education is strongest at Key Stage 4 (Model 7 and 8 in Table A2.1 in the appendix).

**Table 2.2** Percentage of children born in each month by parental education

<b>Millennium Cohort Study (MCS)</b>												
Relative age (Month of birth)												
	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
	(Sep)	(Oct)	(Nov)	(Dec)	(Jan)	(Feb)	(Mar)	(Apr)	(May)	(Jun)	(Jul)	(Aug)
<i>Parental education</i>												
Low	8.69	8.34	8.23	9.31	8.61	7.15	7.98	8.23	8.23	8.92	8.21	8.09
High	8.52	8.39	8.86	8.24	7.85	7.26	9.11	8.02	8.59	8.71	8.27	8.17
<i>N</i> = 9,252												
Pearson $\chi^2(11) = 9.7380$ (Pr = 0.554)												
Cramér's V = 0.0324												
<b>Longitudinal Study of Young People in England (LSYPE)</b>												
Relative age (Month of birth)												
	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
	(Sep)	(Oct)	(Nov)	(Dec)	(Jan)	(Feb)	(Mar)	(Apr)	(May)	(Jun)	(Jul)	(Aug)
<i>Parental education</i>												
Low	7.92	8.34	8.08	7.82	8.09	7.93	8.47	7.93	8.97	8.57	9.08	8.81
High	8.04	8.24	7.72	8.29	7.85	8.29	8.35	8.01	8.75	8.87	8.76	8.81
<i>N</i> = 13,916												
Pearson $\chi^2(11) = 3.4657$ (Pr = 0.983)												
Cramér's V = 0.0158												

*Source:* Millennium Cohort Study (MCS), Longitudinal Study of Young People in England (LSYPE).

**Table 2.3** The effects of relative age at school entry and social origin on cognitive and educational outcomes at different ages

	BAS Vocabulary at age 5 (MCS)		Key Stage 2 score at age 11 (LSYPE)		Key Stage 3 score at age 14 (LSYPE)		5 GCSE at grades A* to C at age 16 (LSYPE)		Academic education after age 16 (LSYPE)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Relative age	-0.035* (0.003)	-0.037* (0.004)	-0.034* (0.002)	-0.038* (0.004)	-0.022* (0.003)	-0.026* (0.004)	-0.006* (0.001)	-0.010* (0.002)	-0.004* (0.001)	-0.006* (0.002)
High parental education	0.545* (0.020)	0.572* (0.038)	0.579* (0.022)	0.621* (0.034)	0.666* (0.023)	0.711* (0.035)	0.272* (0.010)	0.309* (0.017)	0.207* (0.011)	0.232* (0.019)
Relative age X High parental education		0.005 (0.006)		0.008* (0.005)		0.008* (0.005)		0.007* (0.003)		0.005* (0.003)
Male	-0.056* (0.020)	-0.056* (0.020)	-0.070* (0.022)	-0.070* (0.022)	-0.101* (0.024)	-0.101* (0.024)	-0.078* (0.010)	-0.078* (0.010)	-0.021* (0.011)	-0.021* (0.011)
N	9,093	9,093	13,077	13,077	13,032	13,032	13,473	13,473	9,454	9,454

*Note:* Models are estimated using data from the Millennium Cohort Study (MCS) and the Longitudinal Study of Young People in England (LSYPE). (1) to (6) are OLS regression models with standardized outcome variables. (7) to (10) are Linear Probability Models with binary outcome variables. Standard errors in parentheses.

Significance level: \* p < 0.10.

Overall, these findings support the hypothesis that the compensatory advantage develops during the school career and is actually strongest, at the stage of the educational career that is most consequential for final educational attainment, taking the GCSE (or equivalent) exams and deciding whether to continue with an academically oriented education at age 16.

### Testing the Mechanisms Bringing About the Compensatory Advantage

Next, we test three possible mechanisms that could be causing the compensatory advantage of students from highly educated families: private lessons, parental help with homework, and school choice. Estimates in Table 2.4 refer to GCSE results at age 16.

First, we look at the role of private lessons. Taking private lessons has a positive impact on educational performance, but this impact does not vary with the relative age at which a child enters school. Furthermore, once one adds private lessons and the interaction between private lessons and relative age to the model, the coefficient for the interaction between parental education and relative age does not vary. If highly educated parents use private lessons to help their children who were born in August to compensate for their initial disadvantage, we should expect a larger reduction in the size of the interaction effect between relative age and parental education. We conclude that private lessons do not bring about the observed compensatory effect.

Second, we look at the role parental help with homework plays in compensating for a disadvantageous month of birth. Help with homework is positively associated with all four measures of educational performance at age 16. However, since the size of the interaction between relative age and parental education does not decrease once we control for help with homework in the models, we also conclude that help with homework does not underlie the compensatory advantage of children of highly educated parents.

Third, we test for the influence of school choice. Model 7 and Model 8 in Table 2.4 report the results of LPM with and without school-fixed effects. In line with the estimates in Ermisch and Del Bono (2012), the effect of parental education is reduced by more than one third in the school-fixed effects models. This finding suggests that a large part of the observed family background inequality in educational outcomes is mediated by school choice. The coefficient of the relative age effect, however, does not change. This result implies that the disadvantage associated with a young age at school start persists within schools. This means that the processes commonly subsumed under the label of school effects, such as the quality

of the school, the socioeconomic composition of peers, and the quality of teachers apparently do not play a role in reducing the month of birth penalty.

### Robustness Checks

We conducted several robustness checks in order to ensure that the substantive conclusions of this chapter hold under different specifications. At each educational stage we used several different outcome measures: we have two other cognitive tests available at age 5; we have considered the results for Mathematics and English at Key Stages 2, and 3; and at Key Stage 4 we have also used the total number of GCSE with grades A\* to C, standardized with a mean of 0 and a standard deviation of 1, as well as the GCSE new style point score, which we also standardized. The findings of these additional analyses are fully in line with those that we presented above and are reported in Table A2.1 in the appendix. The results of the standardized scores at Key Stage 4 underscore the conclusion that the compensatory effect is largest at Key Stage 4.

Furthermore, we estimated our models only for those students born just before and just after the discontinuity created by the cut-off date for admission into primary school. We construct a dummy that equals 1 for children born between June and August and 0 for children born between September and November. The children born in other months are dropped from the analysis. Table A2.2 reports the results of these models. The results of this alternative specification are fully in line with the results presented above.

Finally, the last robustness check reported in Table A2.3 in the appendix distinguishes between three instead of two levels of parental education: a high parental education group to which those children belong who have at least one parents who attended higher education, a middle education group of those children who have at least one parent with *A Level*, and a group of the children without a parent who has *A Level*. The results show evidence of a compensatory effect concentrated in the high parental education group. This result is, hence, also fully in line with the conclusions of this chapter.

**Table 2.4** The effects of relative age at school entry and social origin on educational outcomes and the mediating role of mechanisms

	5 GCSE at grades A* to C at age 16							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
								With School-Fixed Effects
Relative age	-0.010* (0.002)	-0.010* (0.002)	-0.010* (0.002)	-0.011* (0.002)	-0.011* (0.002)	-0.011* (0.004)	-0.006* (0.001)	-0.006* (0.001)
High parental education	0.311* (0.017)	0.281* (0.017)	0.280* (0.017)	0.307* (0.017)	0.303* (0.017)	0.303* (0.017)	0.272* (0.010)	0.174* (0.008)
Relative age X High parental education	0.007* (0.003)	0.007* (0.003)	0.007* (0.003)	0.008* (0.003)	0.008* (0.003)	0.008* (0.003)		
Male	-0.078* (0.010)	-0.071* (0.010)	-0.071* (0.010)	-0.073* (0.010)	-0.071* (0.010)	-0.071* (0.010)	-0.078* (0.010)	-0.094* (0.008)
Private lessons		0.176* (0.011)	0.184* (0.018)					
Relative age X Private lessons			0.001 (0.003)					
Help with homework					0.063* (0.013)	0.066* (0.025)		
Relative age X Help with Homework						0.000 (0.004)		
N	13,447	13,447	13,447	13,257	13,257	13,257	13,473	13,473

Note: All models are Linear Probability Models and are estimated using data from the Longitudinal Study of Young People in England (LSYPE). Standard errors in parentheses.

Significance level: \* p < 0.10.

## Discussion and Conclusion

In this chapter, we have provided evidence that a young school entry age in England, a country with strict rules for admission to primary education, entails a long-lasting disadvantage for educational attainment. We have also shown, however, that a young age at school start has less harmful consequences for those with highly educated parents. In their case, the month of birth penalty accounts for a reduction of three percentage points in the probability that a child achieves at least five GCSE (and equivalents) with grades between A\* and C, a key achievement for later enrollment at a university. For children of parents with low education, the month of birth penalty rises to eleven percentage points. We interpret these findings as causal evidence of a more general compensatory advantage enjoyed by children of highly educated parents (Bernardi 2014).

In addition, we have shown that the compensatory advantage for those with highly educated parents is not in place before school starts but that it emerges later in the educational career. It actually becomes strongest when the first important transition in the English education system takes place.

These results lead to the question of how the compensatory effect of social origin comes about. Although we have not been able to provide a satisfactory answer to this question, we can at least exclude some of the “usual suspects”. Our findings show that help with homework, contracting private tutors, and school choice do not explain the observed compensatory advantage. However, several other mechanisms that we have not been able to test may be at play. Insights, in this respect, can come from the economic literature on birth endowments and parental responses to them (Almond and Mazumder 2013). In this area, the key question addressed is whether parents reinforce or reduce initial differences in endowments between their children (see also Conley 2004).

Classic work in sociology of education on teachers’ expectations and labeling may also explain how initially disadvantaged children from highly educated families catch up with their peers, while similarly disadvantaged children from low educated families become trapped in a trajectory of low achievement (Hargreaves et al. 1975; Jussim and Harber 2005). Since children from disadvantaged families are more often subject to negative teacher labeling, an initial low achievement linked to month of birth is likely to be more harmful for their future educational outcomes.

Throughout this chapter we have argued that month of birth provides a unique opportunity to test the compensatory effect of social origin hypothesis in those countries with

strict school admission rules. We have also shown that the penalty associated with an unlucky month of birth is comparable in size to the much-discussed gender inequality. In order to test how common the phenomenon of a compensatory effect is, further research could employ causal identification strategies to test whether a compensatory advantage is observed for other individual or family characteristics that are associated with a disadvantage in educational outcomes, such as birth order and parental separation.

## ENDNOTES

1. A sensitivity analysis shows that with our current sample size, fixing alpha at 0.10 we would be able to depict a Cohen's  $W$  equal to 0.04 in the MCS ( $N = 9,252$ ) and equal to 0.03 in the LYSPE ( $N = 13,916$ ) with a power equal to 0.80 (Faul et al. 2009).

## REFERENCES

- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics*, 124, 1729–1772.
- Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annual Review of Economics*, 5, 37–56.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Bedard, K. and Dhuey, E. (2006). The persistence of early childhood maturity: international evidence of long-run age effects. *The Quarterly Journal of Economics*, 121, 1437–1472.
- Becker, B. (2011). Social disparities in children's vocabulary in early childhood: does preschool education help to close the gap? *British Journal of Sociology*, 62, 69–88.
- Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: a regression discontinuity based on month of birth. *Sociology of Education*, 87, 74–88.
- Bernardi, F. and Cebolla-Boado, H. (2014). Previous school results and social background: compensation and imperfect information in educational transitions. *European Sociological Review*, 30, 207–17.
- Black S. E., Devereux, P. J., and Salvanes, K. G. (2011). Too young to leave the nest? The effects of school starting age. *Review of Economics and Statistics*, 93, 455–67.

- Boudon, R. 1998. Social mechanisms without black boxes. In P. Hedström and R. Swedberg (Eds.), *Social Mechanisms: An Analytical Approach to Social Theory* (pp. 172–203). Cambridge: Cambridge University Press.
- Breen, R. and Goldthorpe, J. H. (1997). Explaining educational differentials: towards a formal rational action theory. *Rationality and Society*, 9, 275–305.
- Buchmann, C. and DiPrete, T. A. (2006). The growing female advantage in college completion: the role of family background and academic achievement. *American Sociological Review*, 71, 515–541.
- Buckles, K. and Hungerman, D. M. (2013). Season of birth and later outcomes: old questions, new answers. *Review of Economics and Statistics*, 95, 711–724.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–120.
- Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York: Pantheon Books.
- Crawford, C., Dearden, L., and Meghir, C. (2010). When you are born matters: the impact of date of birth on educational outcomes in England. *IFS Working Papers*, No. 10/06.
- Crawford, C., Dearden, L., and Greaves, E. (2013). The impact of age within academic year on adult outcomes. *IFS Working Papers*, No. 13/07.
- Crawford, C., Dearden, L., and Greaves, E. (2014). The drivers of month of birth differences in children’s cognitive and non-cognitive skills: a regression discontinuity analysis. *Journal of the Royal Statistical Society: Series A*, DOI: 10.1111/rssa.12071.
- Cunha, F. and Heckman, J. J. (2007). The technology of skill formation. *American Economic Review*, 97, 31–47.
- Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*. Amsterdam: Elsevier.
- Department for Education and National Centre for Social Research. (2011). Longitudinal Study of Young People in England: Waves One to Seven, 2004-2010 (computer file). 11<sup>th</sup> Edition. Colchester, Essex: UK Data Archive (distributor).
- Dobkin, C. and Ferreira, F. I. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review*, 29, 40–54.
- Ermisch, J. and Del Bono, E. (2012). Inequality in achievements during adolescence. In J. Ermisch, M. Jäntti and T. Smeeding (Eds.), *From Parents to Children: The*

- Intergenerational Transmission of Advantage* (pp. 285–308). New York, NY: Russel Sage Foundation.
- Eurydice. (2011). *Grade Retention during Compulsory Education in Europe: Regulations and Statistics*. Brussels: Eurydice.
- Faul, F., Erdfelder, E. Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
- Fredriksson, P. and Öckert, B. (2014). Life-cycle effects of age at school start. *The Economic Journal*, 124, 977–1004.
- Freedman, D., Pisani, R., and Purves, R. (2007). *Statistics. Fourth Edition*. New York: W.W. Norton & Company.
- Hargreaves, D., Hester, S., and Mellor, F. J. (1975). *Deviance in Classrooms*. London: Routledge.
- Hsin, A. (2012). Is biology destiny? Birth weight and differential parental treatment. *Demography*, 49, 1385–1404.
- Jussim, L. and Harber, K. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131–155.
- Mood, C. (2010). Logistic regression: why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26, 67–82.
- Mühlenweg, A. M. and Puhani, P. A. (2010). The evolution of the school-entry age effect in a school tracking system. *Journal of Human Resources*, 45, 407–438.
- Restrepo, B. (2012). Who compensates and who reinforces? Parental investment responses to child endowment shocks. *Unpublished Manuscript*, available at [http://www.mwpweb.eu/1/154/resources/research\\_255\\_1.pdf](http://www.mwpweb.eu/1/154/resources/research_255_1.pdf), accessed August 26, 2015.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., and Taggart, B. (2010). *Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education Project*. Oxford: Routledge.
- Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40, 1008–1018.

- Tymms, P., Merrel, C., and Henderson, B. (1997). The first year at school: A quantitative investigation of the attainment and progress of pupils. *Educational Research and Evaluation*, 2, 101–18.
- University of London. Institute of Education. Centre for Longitudinal Studies. (2010). Millennium Cohort Study: Third Survey, 2006 (computer file). 4<sup>th</sup> Edition. Colchester, Essex: UK Data Archive (distributor).

**Table A2.1** The effects of relative age at school entry and social origin on cognitive and educational outcomes at different ages, alternative outcomes

	Scores on the British Ability Scale (BAS) at age 5 (MCS)		KS2 score at age 11 (LSYPE)		KS3 score at age 14 (LSYPE)		Key Stage 4 at age 16 (LSYPE)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Picture similarity	Pattern construction	English	Math	English	Math	Number of GCSE at grades A* to C	GCSE new style point score
Relative age	-0.030* (0.004)	-0.042* (0.004)	-0.036* (0.004)	-0.036* (0.004)	-0.025* (0.004)	-0.027* (0.004)	-0.019* (0.004)	-0.013* (0.004)
High parental education	0.345* (0.042)	0.389* (0.042)	0.597* (0.034)	0.553* (0.034)	0.646* (0.037)	0.655* (0.035)	0.699* (0.034)	0.643* (0.034)
Relative age X High parental education	0.003 (0.006)	0.001 (0.006)	0.009* (0.005)	0.006 (0.005)	0.006 (0.005)	0.009* (0.005)	0.012* (0.005)	0.007* (0.005)
Male	-0.107* (0.022)	-0.177* (0.022)	-0.319* (0.021)	0.103* (0.023)	-0.351* (0.024)	0.035* (0.024)	-0.204* (0.022)	-0.214* (0.024)
N	9,105	9,066	12,968	12,969	12,655	12,841	13,473	13,473

*Note:* Models are estimated using data from the Millennium Cohort Study (MCS) and the Longitudinal Study of Young People in England (LSYPE). All models are OLS regression models with standardized outcome variables. Standard errors in parentheses. Significance level: \* p < 0.10.

**Table A2.2** The effects of relative age at school entry and social origin on cognitive and educational outcomes at different ages, alternative specification

	Key Stage 2 score at age 11 (LSYPE)	Key Stage 3 score at age 14 (LSYPE)	5 GCSE with grades A* to C at age 16 (LSYPE)
	(1)	(2)	(3)
Young at school entry (dummy)	-0.325* (0.032)	-0.211* (0.032)	-0.074* (0.016)
High parental education	0.528* (0.034)	0.603* (0.034)	0.231* (0.016)
Young at school entry (dummy) X High parental education	0.104* (0.047)	0.100* (0.046)	0.043* (0.023)
Male	-0.083* (0.023)	-0.117* (0.023)	-0.103* (0.011)
<i>N</i>	6,615	6,577	6,814

*Note:* Models are estimated using data from the Millennium Cohort Study (MCS) and the Longitudinal Study of Young People in England (LSYPE). (1) and (2) are OLS regression models with standardized outcome variables. (3) is a Linear Probability Models with a binary outcome variable. Standard errors in parentheses.

Significance level: \*  $p < 0.10$ .

**Table A2.3** The effects of relative age at school entry and social origin on cognitive and educational outcomes at different ages, alternative specification of parental education using three categories

	Key Stage 2 score at age 11 (LSYPE)	Key Stage 3 score at age 14 (LSYPE)	5 GCSE with grades A* to C at age 16 (LSYPE)
	(1)	(2)	(3)
Relative age	-0.036* (0.003)	-0.023* (0.003)	-0.009* (0.002)
Middle parental education	0.420* (0.042)	0.432* (0.041)	0.162* (0.020)
High parental education	0.725* (0.035)	0.841* (0.035)	0.334* (0.017)
Relative age X Middle parental education	0.006 (0.007)	0.005 (0.007)	0.001 (0.003)
Relative age X High parental education	0.009* (0.005)	0.008* (0.005)	0.007* (0.003)
Male	-0.078* (0.016)	-0.112* (0.016)	-0.093* (0.008)
<i>N</i>	13,077	13,032	13,473

*Note:* Models are estimated using data from the Longitudinal Study of Young People in England (LSYPE). (1) to (2) are OLS regression models with standardized outcome variables. (3) is a Linear Probability Models with a binary outcome variable. Standard errors in parentheses.

Significance level: \*  $p < 0.10$ .

# **CHAPTER 3: WHEN GROWING UP WITHOUT A PARENT DOES NOT HURT: PARENTAL SEPARATION AND THE COMPENSATORY EFFECT OF SOCIAL ORIGIN<sup>†</sup>**

## **Introduction**

Numerous studies have shown that children who experience parental separation during childhood achieve lower educational outcomes (Amato 2000, 2010; McLanahan et al. 2013). Most of these studies do not take into account the possible heterogeneity of separation effects. It may well be that the effects of parental separation on child outcomes differ across social origin groups and that average effects misrepresent the true consequences of parental separation.

Only few studies look at the variation of the association between parental separation and children's educational outcomes with social origin (Albertini and Dronkers 2009; Augustine 2014; Bernardi and Radl 2014; Bernardi et al. 2014; Fischer 2007; Mandemakers and Kalmijn 2014). These studies do not control for selection into parental separation on unobserved variables. Testing the heterogeneity of parental separation effects employing such a research design, namely, family-fixed effects models, is the central contribution of this chapter to the literature.

On a theoretical level, I develop a precise prediction on the direction of the interaction between social origin and parental separation: the effects of parental separation should be weaker among higher class families. I label this prediction the compensatory class hypothesis and derive it from two sources.

First, it follows from the resource perspective which directly connects parental resources and educational outcomes. This is one of the mechanisms that are supposed to underlie the parental separation penalty (Jonsson and Gähler 1997). It can be expected that the resources of the non-resident parent will have an effect on the educational outcomes of the child even after separation and that the amount of resources available is likely to be a function of family background.

---

<sup>†</sup> A shortened version of this chapter was published in 2015 as "When Growing Up Without a Parent Does Not Hurt: Parental Separation and the Compensatory Effect of Social Origin." *European Sociological Review*, 31, 546-557. DOI: 10.1093/esr/jcv057. The data used in this chapter were made available to me by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin.

Second, the compensatory effect of social origin can be expected based on research on educational inequalities. Parents from different social classes have different amounts of resources available and may act differently in order to compensate for the negative consequences of any disadvantageous characteristic or life event (Bernardi 2014; Conley 2004). Parental separation is one of the major disadvantageous life events which can affect a child's educational career. The analysis of the consequences of parental separation provides hence an ideal case to test the compensatory effect of social origin hypothesis.

## **Background**

### How Parental Separation May Affect Children's Educational Outcomes

There are two main mechanisms through which parental separation may negatively affect children's educational outcomes. First, from a resource perspective, growing up in a family where only one parent is constantly present may be disadvantageous. Parental separation may reduce financial, social, and time resources devoted to the child (Astone and McLanahan 1991; Jonsson and Gähler 1997). Loss of these resources may lead to lower educational outcomes. However, parental separation does not lead to actual changes in resources parents invest in their children if parents provide the same amount of resources to their children before and after separation. The involvement of parents after separation is likely to be a function of social origin. This links the resource perspective and the compensatory class hypothesis which argues that higher class families can compensate the negative effects of parental separation.

Second, in the literature on separation effects there is a considerable emphasis on the possible psychological consequences of experiencing parental break-up. The stress-adjustment perspective argues that a change in family structure induces stress which leads to an under-performance of the child (Amato 2000). If we assume that stress experienced by the child following the separation is the main mechanism bringing about the negative effect of parental separation on child education, the consequences of the separation are initially probably the same for children from different social classes. However, children from higher social origin families may be better able to cope with the stress in the long run. This may allow them to compensate the negative effects of parental separation.

## The Compensatory Effect of Social Origin

With few exceptions, research on the consequences of separation for children's educational outcomes has not focused on the question of whether the consequences of parental separation vary with social origin (Albertini and Dronkers 2009; Augustine 2014; Bernardi and Radl 2014; Bernardi et al. 2014; Fischer 2007; Mandemakers and Kalmijn 2014). Not taking into account the heterogeneity of separation effects, however, may misrepresent the true impact of parental separation on child outcomes (Amato 2010).

Besides this methodological point, there are theoretical reasons to study the heterogeneity in separation effects. From a perspective of research on educational inequalities, it can be expected that the effects of parental separation differ with family background. This follows from what can be labeled the compensatory class hypothesis (Bernardi 2014; Conley 2004). According to this hypothesis the consequences of disadvantageous life events are less harmful for children from higher class families. Examples of studies which demonstrate such a compensatory effect look at differences between siblings (Conley 2004), prenatal exposure to nuclear radiation (Almond et al. 2009), birth weight (Torche and Echevarría 2011), and school entry age (Bernardi 2014).

Certainly, there are differences between these life events and parental separation. However, similar to these examples, parental separation can be viewed as a shock which influences negatively child outcomes. There are three main reasons why a compensatory effect can be expected with respect to parental separation.

First, different classes have different amounts of financial and social resources available, which can be mobilized in order to overcome the negative consequences of the separation (Fischer 2007). For instance, higher class parents can pay for private lessons if the school results of their children become worse following separation. They also have more social resources, including friends and the extended family, allowing them to replace the parent who leaves the household.

Second, higher class families who experience parental separation may see the risk of downward social mobility of their children which can lead them to employ countersteering measures (Breen and Goldthorpe 1997). For instance, if the father has a higher class position than the mother and leaves the household, he can, through keeping close contact with the child, make sure that the child aspires to achieve his level of education and his occupational position. There is empirical evidence which suggests that children in higher class families have more contact with the non-resident parent following parental separation. Arditti and

Keith (1996) show that fathers with a higher socio-economic status visit their children more frequently following divorce. Similarly, Cooksey and Craig (1998) report an increase in contact with the child of the non-resident father with rising level of education. Cheadle et al. (2010) report that children of highly educated mothers are more likely to be in the group of children with separated parents who have regular high contact with their non-resident fathers. In addition, Westphal et al. (2014) show that overnight stays of children with the non-resident father are more likely if he has a high level of education.

Third, higher educated parents may provide a more stable environment following separation (Augustine 2014; Mandemakers and Kalmijn 2014). As a consequence, the lives of upper class children may be less disrupted by parental separation so that the child can faster adapt to the new situation. Such an argument can also receive support by the ethnographic study of Lareau (2011). Lareau (2011) draws a sharp distinction between the childrearing practices of parents from upper and lower class families. The upper class families raising style around organized activities may be less affected by parental separation than the raising style of lower class families since the organized activities will continue following separation.

#### Previous Studies on Social Origin Differences in the Association between Parental Separation and Children's Educational Outcomes

Few studies have analyzed the variation of the association between parental separation and children's educational outcomes with social origin. Jonsson and Gähler (1997), for instance, do not focus on the variation of this association. They, however, argue that if a high resource father leaves the household, this implies an experience of downward social mobility for the child.

Using data on Italy, Albertini and Dronkers (2009) find a negative association between parental separation and child education in families with a low level of maternal education. They find, however, no similar association in families with a highly educated mother.

Fischer (2007) uses data on the Netherlands to analyze the variation of the association between parental divorce and children's education with paternal and maternal education and occupational status. She finds the negative association between parental divorce and child education to be decreasing with higher levels of maternal resources and to be increasing with higher levels of paternal resources (controlling for both in the same models).

Mandemakers and Kalmijn (2014) study the associations between parental union dissolution, parental resources, and children's test scores at age 10 with data from the 1970

British Cohort Study. They find that the negative association between parental separation and reading test scores is weaker in families where the father has a high level of education but find no social origin differences in the association between separation and mathematics test scores.

Bernardi and Radl (2014) use the Generations and Gender Survey to study the associations between parental break-up, education of the parent the child lived with, and the completion of tertiary education of children in 14 countries (Germany is not among them). They report that the negative association between parental break-up and tertiary education is stronger if the child lived with a low educated parent in countries with early selection in the education system and is stronger if the child lived with a highly educated parent in countries with late selection.

Augustine (2014) tests whether the association between parental separation and child outcomes varies with maternal education in the United States. She uses data from the NICHD Study of Early Child Care and Youth Development (United States) and finds that a family structure without both parents present in the household is connected to lower parenting quality only in families of less educated mothers.

Bernardi et al. (2014) use the 1970 BCS, as Mandemakers and Kalmijn (2014), whilst employing university attendance, access to the service class, and occupational position as outcome variables. They argue that a parental divorce after age 5 is associated with a lower probability of attending university and having a service class job for all children, however, the negative association is stronger if both parents have themselves university education compared with if both parents have a high school degree or less. They find no differences between social origin groups for occupational status.

To sum up, findings from these six studies are mixed. The divergence in results can be due to differences in analyzed outcomes, in the operationalization of parental separation and social origin, or in analyzed countries and cohorts. However, to my mind the most important shortcoming of previous studies is that they do not control for selection into parental separation on unobserved variables. The interaction effects reported in these studies may be spurious due to differences between social origin groups in selection into parental separation.

## Data and Methods

### Research Design

A major problem in the study of the consequences of parental separation is that unobserved variables may confound the relationship between parental separation and child outcomes (Ní Bhrolcháin 2001). People self-select into separation and the same characteristics which make parents more likely to separate may lead to lower child outcomes. Cherlin et al. (1991) demonstrate that taking into account some measurable pre-divorce characteristics of the parents considerably lowers the estimated differences in test scores between children from divorced and non-divorced families. Because it does not seem possible to measure all parental characteristics that can lead to selection into separation, a research design that controls for selection into parental separation on unobserved variables supports stronger causal interpretations of separation effects. There are many variables that are hard to measure but may influence both parental separation and child outcomes. For instance, parents who separate may be less interested in providing their children with education than parents who do not separate.

During the past decades different approaches have been proposed to overcome this identification problem (Amato 2000; McLanahan et al. 2013). Good instruments are rare and instrumental variable (IV) estimation strategies based on changes in divorce laws have to deal with the problem that law changes may affect child outcomes in other ways than only via increasing the number of divorces (Corak 2001; Gruber 2004). Furthermore, such an approach cannot take into account the dissolution of non-marital unions which are becoming more important with growing cohabitation. An alternative is to use parental death as a quasi-experiment for parental loss (Corak 2001; Francesconi et al. 2010). However, mortality is related to confounding characteristics and, thus, parental death is not randomly distributed across the population (see e.g. Torssander and Erikson 2010). This should disqualify the use of parental death as a quasi-experiment for parental loss.

An alternative to control for selection into parental separation are fixed effects models. Individual-fixed effects models have been used in which the same children are compared before and after they experienced parental separation (Aughinbaugh et al. 2005). Another type of fixed effects models applies fixed effects at the family level (Björklund and Sundström 2006; Björklund et al. 2007; Ermisch and Francesconi 2001; Francesconi et al. 2010; Sandefur and Wells 1999; Sigle-Rushton et al. 2014). These so-called family-fixed effects (or

sibling difference) models use only variation between siblings to identify the effects of separation. The advantage of this method is that it controls for unobserved confounding variables, which are constant between siblings at the family level and in the environment, in particular characteristics of the parents, which are shared among siblings. Most of the confounding variables discussed in the separation literature are at the parental level which makes this a convincing research design to control for selection into parental separation. For that reason, I apply family-fixed effects models throughout this chapter.

The limitations of family-fixed effects models have to be kept in mind when interpreting the results presented below. Family-fixed effects models make it necessary to select a sample of siblings. Hence, children from one-child families cannot be included in these models (McLanahan et al. 2013). It is difficult to test whether the effects of parental separation differ between children in those families and children who grow up with siblings without another research design which allows researchers to identify causal effects. However, below I compare descriptive statistics and naive regression results between the full and the siblings sample. These additional analyses do not provide any indication that the siblings sample differs from the full sample based on observed variables.

More generally, Frisell et al. (2012) show that family-fixed effects models can lead to biased estimates if confounders are not completely shared among siblings. In the case of parental separation, the main unobserved variables which lead to selection into separation are at the parental level and, hence, shared among siblings. Even if there are some unshared confounders, their role should be smaller than the role of shared confounders. For that reason family-fixed effects estimates are less biased than naive estimates.

Family-fixed effects models, however, do not control for reverse causality that is if children's characteristics, e.g. cognitive problems, influence parent's decision to separate. Furthermore, it may be that siblings are not influenced in the same way by parental characteristics (Ermisch and Francesconi 2001). Family-fixed effects models do not control for selection on parental characteristics which are not fully shared among siblings.

Finally, family-fixed effects models compare the experience of a change in household family structure between siblings. They do not, however, compare differences in the experience of parental conflict. If there is parental conflict already before a change in household structure, this is not taken into account. Therefore the possible consequences of parental separation may be underestimated. However, I do not expect that this should bias the estimation of social origin differences in separation effects, which is the main focus of this chapter.

Another drawback of the fixed effects model is that the information used to identify the effects of parental separation is reduced to those siblings who differ in their experience of parental separation. Because their number is small, this leads to rather imprecise estimates.

## Data

I use data from version 28 of the German Socio-Economic Panel Study (SOEP) which is a representative household panel study including information on education, occupation, health, and other demographic topics (Wagner et al. 2007). Version 28 includes yearly collected data from 1984 till 2011.

The sample is restricted to those respondents who filled out the youth questionnaire between 2000 (the year this questionnaire was introduced) and 2011. These data are collected annually through face-to-face interviews and information is provided by all children who grow up in SOEP households in the year in which they turn 17 years (Giesselmann and Staneva 2012). The information refers to the most recent educational outcomes at age 16 years. The response rate of the youth questionnaire was 87.2% in 2010 and 84.2% in 2011 (TNS Infratest Sozialforschung 2012).

In the family-fixed effects models I restrict the sample to all respondents who have another sibling who has also filled out the questionnaire. I dropped children who experienced a death of a parent from the analysis sample. The final sample includes 1,947 children from 904 families who were born between 1982 and 1994.

Since nearly all families send their children to schools in the same states, the family-fixed effects models control for differences between states. Furthermore, I conducted the analysis for the whole of Germany because all children in the sample experienced most of their childhood in reunified Germany. As a robustness check, I report results restricted to families with an origin in the Federal Republic of Germany (FRG). The robustness check suggests that the compensatory effect of social origin is slightly stronger in this reduced than in the full sample which includes children from families with an origin in the German Democratic Republic (GDR). This result is in line with previous research on educational inequalities in East and West Germany which argued that social origin differences in educational outcomes are smaller in East than in West Germany in the period after German reunification (Kesler 2003).

## Variables and Descriptive Statistics

Parental separation is measured based on whether a child lived all his childhood with both his parents. I use a household-based definition of parental separation because it includes both married and cohabiting couples. Parental separation is assumed to have occurred if a child lives no longer with both parents. I employ two age thresholds. First, in the models on grades at age 16 years, I measure parental separation based on the first 15 years of childhood. The parental separation variable is coded 1 if a child did not live all 15 years of childhood continuously with both parents. Hence, this dummy variable is coded 0 for all children who lived continuously with both parents during 15 years of their childhood. The age threshold of 15 years is taken because educational outcomes refer to age 16 years and it ensures that the separation of the parents happened before the outcome is measured. With respect to track attendance as an outcome variable, I only take into account 11 years of childhood. This is done because the crucial allocation to tracks happens for most children in the eleventh year of childhood. When a child experiences that a parent leaves the household for longer time this counts as an experience of parental separation, even if the parent later returns into the same household. I prefer to code parental separation this way because both the mechanisms which underlie the parental separation penalty and the mechanisms which can bring about the compensatory class effect can operate in such a situation.

I measure social origin via parental education defined as the highest level of education of any parent. I employ a dummy variable, which is coded 1 if at least one parent has *Abitur* (the degree obtained upon completion of the upper track in the German secondary school system and a requirement for university entry, comparable to *A Levels* in the UK) or an equivalent qualification and 0 otherwise. In one part of the analysis, I look at father's and mother's education separately. Parental education as a measure of social origin has the advantage that the educational levels of the parents are largely stable over their life courses. In particular the levels of education of the parents are not changing after parental separation, as this may happen to their occupational attainments.

I analyze the effects of parental separation on educational outcomes in Germany (Bohrhardt 2000; Francesconi et al. 2010). Germany has a tracked school system which allocates students in most states to three tracks in secondary school after four years in primary school. Students have to complete the highest track (*Gymnasium*) in order to attend university. In some states the two lower tracks are combined into one track.

Track allocation takes place around age 10. This placement process is crucial because changes from one track to another at a later point in a child's school career are, despite being possible, rather uncommon. Consequently, research on educational transitions in Germany has demonstrated that the initial track attended is an important predictor of final educational attainment (Hillmert and Jacob 2010).

I look at three educational outcomes at age 16 years: the school grade in Mathematics, the school grade in German, and the probability of attending the highest track. In the German school system, grades from 1 to 6 are given. I rescale the grades so that a higher grade signifies a better performance. The grades are estimated using Ordinary Least Squares (OLS) regression models with robust standard errors.

The attendance of the upper track is a dichotomous variable. I compare the attendance of the upper track (*Gymnasium*) to any lower track or early dropout. Respondents who attend comprehensive schools are not used in the models that estimate track attendance because comprehensive schools comprise all tracks. Because comprehensive schools are not very common in Germany, this is only a small number of respondents (122 children). All other outcomes are analyzed including those children.

Estimates on track attendance are calculated using Linear Probability Models with robust standard errors. I use LPM because of the straightforward interpretation of the coefficients, especially with regard to the interaction effects and the possibility to compare coefficients across models (Angrist and Pischke 2009; Mood 2010). The out-of-sample predictions of these models are very low (between 0 and 0.3 percent).

All fixed effects models are estimated using the `xtreg` command in STATA 13.1. Because I have clear hypotheses on the direction of both the separation and the compensatory effect I use one-tailed significance tests in all models (Freedman et al. 2007).

Controls include gender and birth order. Gender controls for male disadvantage. Birth order is an important control because later born children are more likely to experience parental separation and birth order has an effect on educational outcomes in Germany (Härkönen 2014; Sigle-Rushton et al. 2014). In addition, I control for birth year via a set of dummy variables. In the models on grades, I also control for track attendance because the value of grades differs between tracks.

Table 3.1 summarizes descriptive statistics on the siblings sample used in the analysis. About 14 percent of children experienced parental separation within 11 years of childhood. About 19 percent experienced parental separation during 15 years of childhood. I report the standard deviation within families because this variation is used to identify the effects in the

family-fixed effects models. One concern may be that there is not enough variation within families, which allows me to identify effects. The within-family standard deviation reported in Table 3.1 shows that this concern is misplaced.

**Table 3.1** Descriptive statistics for variables in the siblings sample used in the analysis ( $N = 1,947$ )

Variable	Mean	Std. dev.	Std. dev. between families	Std. dev. within families	Min.	Max.	$N$
Age in 2011	23.01	3.28	2.73	1.92	17	29	1,947
Male	0.49	0.50	0.35	0.36	0	1	1,947
Parental separation within 11 years of childhood	0.14	0.35	0.33	0.13	0	1	1,947
Parental separation within 15 years of childhood	0.19	0.40	0.37	0.14	0	1	1,947
Birth order	1.90	0.97	0.69	0.64	1	9	1,947
High parental education	0.34	0.47	- <sup>c</sup>	- <sup>c</sup>	0	1	1,947
High father's education	0.26	0.44	- <sup>c</sup>	- <sup>c</sup>	0	1	1,909
High mother's education	0.23	0.42	- <sup>c</sup>	- <sup>c</sup>	0	1	1,936
GDR origin <sup>a</sup>	0.23	0.42	- <sup>c</sup>	- <sup>c</sup>	0	1	1,934
Mathematics grade	4.06	1.04	0.81	0.67	1	6	1,917
German grade	4.08	0.84	0.65	0.56	1	6	1,922
Upper track attendance ( <i>Gymnasium</i> ) <sup>b</sup>	0.39	0.49	0.43	0.24	0	1	1,825

*Note:* The sample includes all respondents with valid information on at least one other sibling.

<sup>a</sup> 13 respondents who lived neither in the Federal Republic of Germany (FRG) nor in the German Democratic Republic (GDR) in 1989 were assigned a missing value on this variable.

<sup>b</sup> 122 respondents who attend comprehensive schools were assigned a missing value on this variable (see text for explanation).

<sup>c</sup> Variables at the family level that cannot vary between siblings.

*Source:* German Socio-Economic Panel Study (SOEP), v28.

The descriptive statistics show that there are missing values on the dependent variables. von Hippel (2007) recommends not to regress on imputed values of dependent variables. For that reason I do not apply multiple imputation. However, I also do not apply listwise deletion. Rather I prefer to estimate each outcome separately excluding only those cases with missing values on the specific outcome. This strategy maximizes the use of the available information.

Another concern with the approach taken in this paper may be that it is unclear how the results are generalizable to children from one-child families (McLanahan et al. 2013). It is not possible to test whether the effects of parental separation differ between children from one-child families and children who grow up with siblings since only the latter are included in the family-fixed effects models. What is more, only those siblings who diverge in their experience of the timing of parental separation identify the effect of parental separation.

I cannot test whether the consequences of separation differ between these groups. However, I can provide descriptive statistics on observed variables in order to see whether these indicate any crucial differences. Table 3.2 compares three different samples and presents descriptive statistics on all children who experienced parental separation (Sample 1, including those not included in the siblings sample), siblings who do not differ in their experience of the timing of parental separation (Sample 2), and siblings who diverge in their experience of the timing of parental separation (Sample 3).

**Table 3.2** Descriptive statistics for different types of children from different types of families who experienced parental separation

	<b>Sample 1: All children who experienced parental separation<sup>a</sup></b>		<b>Sample 2: Siblings who share the experience of parental separation</b>		<b>Sample 3: Siblings who diverge in their experience of parental separation<sup>b</sup></b>	
	Within 11 years of childhood	Within 15 years of childhood	Within 11 years of childhood	Within 15 years of childhood	Within 11 years of childhood	Within 15 years of childhood
Age in 2011	22.76 (3.39)	22.80 (3.39)	23.18 (3.06)	23.01 (3.06)	22.93 (3.15)	23.21 (3.26)
Male	0.50 (0.50)	0.50 (0.50)	0.46 (0.50)	0.47 (0.50)	0.50 (0.50)	0.50 (0.50)
Birth order	1.79 (0.92)	1.78 (0.90)	1.90 (0.99)	1.90 (0.97)	2.08 (1.00)	2.01 (0.92)
High parental education	0.25 (0.43)	0.29 (0.45)	0.25 (0.43)	0.31 (0.46)	0.27 (0.45)	0.29 (0.46)
<i>N</i>	512	666	208	297	143	165

*Note:* Standard deviation in parentheses.

<sup>a</sup> These comparison groups include children from one-child families and children with no information on any siblings in the data.

<sup>b</sup> At least one sibling experienced parental separation within 11 or 15 years of childhood, at least one sibling did not.

*Source:* German Socio-Economic Panel Study (SOEP), v28.

There are no obvious indications of crucial differences in descriptive statistics between these three groups, independent of whether 11 or 15 years of childhood are taken into consideration. Birth order is indeed slightly higher in the group of children who diverge in their experience of the timing of parental separation which is as expected because higher order births are more likely to experience parental separation. In any case, I control for birth order effects in the family-fixed effects models (Sigle-Rushton et al. 2014).

But with respect to all other observed variables, these comparisons suggest that differences between these groups can be neglected. For instance, 27-29% of the children who

diverge in their experience of parental separation are from a family with a high level of parental education compared to 25-29% of all children who experience parental separation. As said above, the three groups can still differ on unobserved characteristics so that full confidence in the external validity of the findings cannot be provided. This limitation of the family-fixed effects models should be kept in mind when interpreting the results.

## Results

### The Influence of Parental Separation on Educational Outcomes

Table 3.3 provides both estimates of average effects of parental separation on educational outcomes as well as effect estimates which take into account the interaction between social origin and parental separation.

**Table 3.3** Results of family-fixed effects models predicting the impact of parental separation on educational outcomes

	Mathematics		German		Upper track attendance (Gymnasium)	
	(1)	(2)	(3)	(4)	(5)	(6)
Parental separation <sup>a</sup>	-0.14 (0.13)	-0.24† (0.17)	-0.04 (0.10)	-0.13 (0.13)	-0.05 (0.06)	-0.10* (0.06)
High parental education X Parental separation <sup>a</sup>		0.35† (0.27)		0.29† (0.22)		0.19† (0.14)
Male	0.08† (0.06)	0.08† (0.06)	-0.41** (0.05)	-0.41** (0.05)	-0.06** (0.02)	-0.06** (0.02)
Birth order	-0.18** (0.07)	-0.18** (0.07)	-0.05 (0.05)	-0.05 (0.05)	-0.08** (0.02)	-0.08** (0.02)
Controls for birth year	Yes	Yes	Yes	Yes	Yes	Yes
Controls for track attendance	Yes	Yes	Yes	Yes	No	No
<i>N</i>	1,917	1,917	1,922	1,922	1,825	1,825

*Note:* Robust standard errors in parentheses.

Significance levels: †  $p < 0.10$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

<sup>a</sup> Parental separation refers to 15 years of childhood in the models which predict school grades and to 11 years of childhood in the track attendance models (see text).

*Source:* German Socio-Economic Panel Study (SOEP), v28.

The results in the specification without interaction effects suggest rather low effects of parental separation on educational outcomes in line with previous research applying family-fixed effects models to German data (Francesconi et al. 2010). None of the effects is statistically significant. The model on track attendance predicts, however, a five percentage

points lower probability of attending the upper track for children who experienced parental separation. This seems to be a substantively large effect which is, however, not statistically significant.

One possibility for the rather weak effects of parental separation on education may be that the baseline models do not take into account the heterogeneity of separation effects. The other models reported in Table 3.3 include the interaction between parental separation and parental education testing the compensatory class hypothesis. Note that there is no main effect for parental education included in the models since parental education does not vary between siblings. The interaction between parental separation and parental education reports, hence, the difference in separation effects between social origin groups.

In these models, I do find a strong heterogeneity of separation effects with respect to social origin in the direction predicted by the compensatory effect of social origin hypothesis. All models suggest that the parental separation penalty is stronger for children from low educated families than for children from highly educated families.

Admittedly, the statistical significance of the results is low which may be due to the small number of siblings who differ in their experience of parental separation. However, with respect to all three outcomes, the interaction effect between parental separation and social origin is statistically significant at the ten percent level. Put in other words, we can be certain at a ten percent significance level that the differences in parental separation effects between social origin groups are not due to sampling error. Confidence in the results is increased because the finding of a compensatory effect of social origin is constantly reproduced across all models. As a robustness check, which is reported below, I repeated the analysis only for West German families which strengthened the results in terms of statistical significance.

In order to interpret the sizes of the separation effects, the coefficients have to be analyzed (Firebaugh 2008; Freedman et al. 2007). All three models suggest substantive large effect sizes. The probability of attending the highest track is reduced by 10 percentage points for children from families with a low level of parental education if they experience parental separation. This effect is stronger than the disadvantage which male children face as well as the disadvantage a second born child experiences compared to his first born sibling.

In the case of all three outcome variables the size of the interaction effect between parental separation and parental education is so large that the negative effect of parental separation only persists for families with a low level of parental education. These models predict that children from higher class families are not less likely to attend the upper track if their parents separate than children from higher class families whose parents do not.

The effect of parental separation on track attendance is particularly interesting, as this is arguably the most important outcome in the German education system and an important predictor of final educational attainment. But also the effect of parental separation on school grades is substantial. Children from families with a low level of parental education have, on average, a one quarter point lower grade in Mathematics if they experienced parental separation. Again, this effect is larger than the difference between the first and the second born sibling. In contrast, children from highly educated families do not have lower grades if they experienced parental separation.

### Comparison between Family-Fixed Effects and Naive Regression Estimates

My results differ from some of the results in previous research. This leads to the question whether this is owing to the family-fixed effects models controlling for selection on unobserved variables.

To test the influence of unobserved variables, I compare the family-fixed estimates to estimates obtained using naive regression models in Table 3.4. I report naive regression estimates without fixed effects for both the siblings sample and the full sample, which also includes children without valid information on a sibling. The comparison of the naive regression estimates with the family-fixed effects estimates demonstrates the importance of controlling for unobserved variables. Based on the naive OLS regression estimates I would not have come to the conclusion that the effects of parental separation are less negative for families with a high level of parental education.

Finally, Table 3.4 can also be used as speaking to the question whether the way I selected a sample of siblings leads to sample selection bias. This can be done by comparing the naive regression estimates of the siblings with the estimates of the full sample. The estimates of these two samples are very close to each other. Effects always point in the same direction and effect sizes are very close. This provides further support to the view that the different results which I report applying the family-fixed effects models are owing to these models controlling for unobserved variables and not owing to issues of these models being based on an unrepresentative sample.

**Table 3.4** Results of models without and with family-fixed effects predicting the impact of parental separation on educational outcomes

	Mathematics			German			Upper track attendance ( <i>Gymnasium</i> )		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Naive OLS, full sample <sup>a</sup>	Naive OLS, siblings sample	Family- fixed effects, siblings sample	Naive OLS, full sample <sup>a</sup>	Naive OLS, siblings sample	Family- fixed effects, siblings sample	Naive LPM, Full sample <sup>a</sup>	Naive LPM, siblings sample	Family- fixed effects LPM, siblings sample
Parental separation <sup>b</sup>	-0.11* (0.05)	-0.10† (0.07)	-0.24† (0.17)	-0.07† (0.04)	-0.12* (0.06)	-0.13 (0.13)	-0.12** (0.02)	-0.13** (0.03)	-0.10* (0.06)
High parental education X									
Parental separation <sup>b</sup>	-0.16† (0.10)	-0.04 (0.13)	0.35† (0.27)	-0.03 (0.08)	0.03 (0.10)	0.29† (0.22)	-0.08† (0.05)	-0.04 (0.07)	0.19† (0.14)
Male	0.09** (0.04)	0.06† (0.05)	0.08† (0.06)	-0.38** (0.03)	-0.38** (0.04)	-0.41** (0.05)	-0.07** (0.02)	-0.05** (0.02)	-0.06** (0.02)
Birth order	-0.02 (0.02)	-0.07** (0.02)	-0.18** (0.07)	-0.02† (0.01)	-0.03* (0.02)	-0.05 (0.05)	-0.05** (0.01)	-0.06** (0.01)	-0.08** (0.02)
High parental education <sup>c</sup>	0.09* (0.05)	0.11* (0.06)	0.13* (0.04)	0.13* (0.04)	0.11* (0.05)	0.11* (0.05)	0.46** (0.02)	0.44** (0.02)	
Controls for birth year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for track attendance	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
N	3,103	1,917	1,917	3,110	1,922	1,922	2,975	1,825	1,825

Note: Robust standard errors in parentheses.

Significance levels: † p < 0.10; \* p < 0.05; \*\* p < 0.01.

<sup>a</sup> These comparison groups include children from one child families and children with no information on any siblings in the data.

<sup>b</sup> Parental separation refers to 15 years of childhood in the models which predict school grades and to 11 years of childhood in the track attendance models (see text).

<sup>c</sup> Variable that does not vary between siblings. Therefore, there is no main effect in the fixed effects models.

Source: German Socio-Economic Panel Study (SOEP), v28.

## Parental, Paternal, or Maternal Effects?

The main finding of this chapter is a compensatory class effect not only but also with respect to track attendance which has long-run implications for final educational attainment. This finding begs the question of how this compensation is brought about. A first step in this respect is to analyze whether the compensatory effect mainly occurs through the education of the father or through the education of the mother (Fischer 2007; Mandemakers and Kalmijn 2014). One reason for doing this can be that children have different relationships to their mother and to their father following separation (Fischer 2007).

Specific hypotheses can be formulated why father's or mother's education should be more important following separation. On the one hand, following separation, mother's resources may be more important than father's resources because the child most often stays with the mother (Fischer 2007; Mandemakers and Kalmijn 2014). On the other hand, the arguments in favor of the compensatory class effect which I developed above suggest rather the opposite: the education of the non-resident parent, most of the time the father, should be more important because his investments into the child following separation may be a function of social origin.

**Table 3.5** Results of family-fixed effects models predicting the impact of parental separation on attendance of the upper track (*Gymnasium*) by father's and mother's education

	<b>Upper track attendance (<i>Gymnasium</i>)</b>			
	(1) Highest level of education of either parent	(2) Education of the father	(3) Education of the mother	(4) Education of the father and education of the mother
Parental separation <sup>a</sup>	-0.10* (0.06)	-0.10* (0.06)	-0.09† (0.06)	-0.10* (0.06)
High parental education X Parental separation <sup>a</sup>	0.19† (0.14)			
High father's education X Parental separation <sup>a</sup>		0.23† (0.17)		0.19 (0.20)
High mother's education X Parental separation <sup>a</sup>			0.18 (0.15)	0.06 (0.17)
Controls for male and birth order	Yes	Yes	Yes	Yes
Controls for birth year	Yes	Yes	Yes	Yes
<i>N</i>	1,825	1,790	1,816	1,781

*Note:* Robust standard errors in parentheses.

Significance levels: †  $p < 0.10$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

<sup>a</sup> Parental separation refers to 11 years of childhood (see text).

*Source:* German Socio-Economic Panel Study (SOEP), v28.

In order to test these two opposing hypotheses I have broken down parental education into two separate measures for father's and for mother's education. Results are reported in Table 3.5. They demonstrate that it is mainly the education of the father which brings about the compensatory class effect. A high level of education of the father provides compensation for the negative consequences of parental separation. This is likely to be the case because the father continues to be involved in the education of the child after having moved to another household and more resources help him to be more involved in his child's life. However, there is also a positive interaction effect between parental separation and maternal education suggesting that the education of the mother may as well play a role in the compensation.

I also entered both education of the father and education of the mother simultaneously into the same model. Such a model is difficult to interpret because of the high collinearity between maternal and paternal education (80 per cent of the children have parents with the same level of education). Furthermore, it is unclear in which direction the causal effect between maternal and paternal education is going. Hence, I would interpret this last model with particular caution. But it also suggests that the education of the father may be more important in bringing about the compensatory effect. This conclusion is in line with results for reading scores reported by Mandemakers and Kalmijn (2014).

### Robustness Check

Previous models include families from East and West Germany. Social origin effects, however, may be different between these families. A high level of education from East Germany may be less advantageous than a high level of education from West Germany since these educational credentials were collected in the German Democratic Republic (GDR). Kesler (2003) provides empirical evidence that social origin differences in educational outcomes are smaller in East compared to West Germany for the period after German reunification.

I conducted a robustness check restricting the sample to children whose parents lived in the Federal Republic of Germany (FRG) in 1989. These models exclude all East German families which are defined as families who lived in the German Democratic Republic (GDR) in 1989. A separate analysis of the GDR origin sample led to very imprecise results because of the small sample size and I abstain from reporting and interpreting these models.

The results for reduced sample, i.e. the sample without the GDR families, are reported in Table 3.6. They suggest that the compensatory class effect is stronger for those children

than for children from the full sample which includes children with a GDR origin. Two out of three interactions between social origin and parental separation increase in effect size by excluding children whose parents lived in the GDR in 1989. The parental separation penalty for children with low educated parents is about 16 percentage points with respect to track attendance. In addition, the statistical significance of the interaction effects in these models is greater than in the models using the pooled sample. This shows that within the reduced sample the substantive conclusions of this paper hold to more certainty.

**Table 3.6** Results of family-fixed effects models of the impact of parental separation on educational outcomes in a sample which excludes children whose parents lived in the GDR in 1989

	Mathematics		German		Upper track attendance (Gymnasium)	
	(1)	(2)	(3)	(4)	(5)	(6)
Parental separation <sup>a</sup>	-0.15 (0.16)	-0.33* (0.19)	-0.06 (0.12)	-0.10 (0.15)	-0.09† (0.06)	-0.16** (0.06)
High parental education X Parental separation <sup>a</sup>		0.65* (0.34)		0.15 (0.25)		0.29* (0.16)
Male	0.14* (0.06)	0.14* (0.06)	-0.43** (0.05)	-0.43** (0.05)	-0.05* (0.02)	-0.05* (0.02)
Birth order	-0.20** (0.08)	-0.20** (0.08)	-0.05 (0.06)	-0.05 (0.06)	-0.09** (0.03)	-0.09** (0.03)
Controls for birth year	Yes	Yes	Yes	Yes	Yes	Yes
Controls for track attendance	Yes	Yes	Yes	Yes	No	No
<i>N</i>	1,458	1,458	1,463	1,463	1,385	1,385

*Note:* Robust standard errors in parentheses.

Significance levels: †  $p < 0.10$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

<sup>a</sup> Parental separation refers to 15 years of childhood in the models which predict school grades and to 11 years of childhood in the track attendance models (see text).

*Source:* German Socio-Economic Panel Study (SOEP), v28.

## Discussion and Conclusion

This chapter provides evidence for social origin differences in the effect of parental separation on child education. The negative effect of parental separation is concentrated among children with low educated parents. Families with highly educated parents are able to make their children avoid the negative effects of parental break-up.

The heterogeneity in separation effects reported in other studies may be spurious because these studies do not control for selection into parental separation (Albertini and Dronkers 2009; Augustine 2014; Bernardi and Radl 2014; Bernardi et al. 2014; Fischer 2007; Mandemakers and Kalmijn 2014). My results also overturn previous studies which argue that

parental separation does not affect children's educational outcomes in Germany (Bohrhardt 2000; Francesconi et al. 2010). This conclusion does not hold up if the heterogeneity of separation effects is taken into account.

The limitations of the present study have to be well understood. The statistical significance of the findings is rather low. This is due to the research design I have applied and the small number of children who identify the separation effect. However, the effect sizes are substantial. Furthermore, the external validity of the findings may be questioned because estimates are only based on a sample of siblings. I have, however, not found any indication that this group is unrepresentative for the whole population. Still I cannot rule out the possibility that the effects of separation on child education are different in one-child families. Further research may increase the confidence in the external validity of the findings. This includes replication of the same approach for different countries and cohorts as well as the use of other research designs, for instance IV approaches. The following chapter uses such an IV approach to analyze parental separation effects, including the heterogeneity of these effects by social origin.

In addition, the research presented here could be expanded in two directions. First, the analysis of separation effects separately by mother's and father's education is only a first step in testing the mechanisms that underlie the compensatory class effect. Unfortunately I am not able to test directly the underlying mechanisms with the data used in this chapter because I only have measures of parental behavior at age 16 to 17 years, which refers to a point later in time than the educational outcomes are measured. Testing the underlying mechanisms of the compensatory class effect requires a data set with information on parental involvement measured after separation but before the educational outcomes.

Second, this study adds parental separation to a number of disadvantageous life events which have stronger negative consequences for children coming from lower class families (Almond et al. 2009; Bernardi 2014; Conley 2004; Torche and Echevarría 2011). Further research may test the compensatory class hypothesis with respect to other life events with negative consequences for educational outcomes.

## REFERENCES

Albertini, M. and Dronkers, J. (2009). Effects of divorce on children's educational attainment in a Mediterranean and catholic society. *European Societies*, 11, 137–159.

- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124, 1729–1772.
- Amato, P. R. (2000). The consequences of divorce for adults and children. *Journal of Marriage and Family*, 62, 1269–1287.
- Amato, P. R. (2010). Research on divorce: Continuing trends and new developments. *Journal of Marriage and Family*, 72, 650–666.
- Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Arditti, J. A. and Keith, T. Z. (1993). Visitation frequency, child support payment, and the father-child relationship postdivorce. *Journal of Marriage and Family*, 55, 699–712.
- Astone, N. and McLanahan, S. (1991). Family structure, parental practices, and high school completion. *American Sociological Review*, 56, 309–320.
- Aughinbaugh, A., Pierret, C., and Rothstein, D. (2005). The impact of family structure transitions on youth achievement: Evidence from the children of the NLSY79. *Demography*, 42, 447–468.
- Augustine, J. M. (2014). Maternal education and the unequal significance of family structure for children's early achievement. *Social Forces*, 93, 687–718.
- Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: A regression discontinuity based on month of birth. *Sociology of Education*, 87, 74–88.
- Bernardi, F., Boertien, D., and Popova, D. (2014). Differential effects of parental separation on child outcomes: Are children from higher social backgrounds affected more? *EUI Working Papers Max Weber Programme*, 2014/06.
- Bernardi, F. and Radl, J. (2014). The long-term consequences of parental divorce for children's educational attainment. *Demographic Research*, 30, 1653–1680.
- Björklund A., Ginther, D., and Sundström, M. (2007). Family structure and child outcomes in the USA and Sweden. *Journal of Population Economics*, 20, 183–201.
- Björklund, A. and Sundström, M. (2006). Parental separation and children's educational attainment: A siblings analysis on Swedish register data. *Economica*, 73, 605–624.
- Borhardt, R. (2000). Familienstruktur und Bildungserfolg: Stimmen die alten Bilder? *Zeitschrift für Erziehungswissenschaft*, 3, 189–207.
- Breen R. and Goldhorpe, J. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, 9, 275–305.

- Cheadle, J. E., Amato, P. R., and King, V. (2010). Patterns of nonresident father contact. *Demography*, 47, 205–225.
- Cherlin, A., Furstenberg, F., Chase-Lansdale, P., Kiernan, K., Robins, P., Morrison, R., and Teitler, J. (1991). Longitudinal studies of effects of divorce on children in Great Britain and the United States. *Science*, 252, 1386–1389.
- Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York, NY: Pantheon Books.
- Cooksey, E. and Craig, P. (1998). Parenting from a distance: The effects of paternal characteristics on contact between nonresidential fathers and their children. *Demography*, 35, 187–200.
- Corak, M. (2001). Death and divorce: The long-term consequences of parental loss on adolescents. *Journal of Labor Economics*, 19, 682–715.
- Ermisch, J. and Francesconi, M. (2001). Family structure and children's achievement. *Journal of Population Economics*, 14, 249–270.
- Firebaugh, G. (2008). *Seven Rules for Social Research*. Princeton, NJ: Princeton University Press.
- Fischer, T. (2007). Parental divorce and children's socio-economic success: Conditional effects of parental resources prior to divorce, and gender of the child. *Sociology*, 41, 475–495.
- Francesconi, M., Jenkins, S., and Siedler, T. (2010). Childhood family structure and schooling outcomes: Evidence for Germany. *Journal of Population Economics*, 23, 1073–1103.
- Freedman, D., Pisani, R., and Purves, R. (2007). *Statistics*. New York: W.W. Norton & Company.
- Frisell, T., Öberg, S., Kuja-Halkola, R., and Sjölander, R. (2012). Sibling comparison designs: Bias from non-shared confounders and measurement error. *Epidemiology*, 23, 713–720.
- Giesselmann, M. and Staneva, M. (2012). BIOAGE17: The Youth Questionnaire. In Goebel, J. (Ed.), *Biography and Life History Data in the German Socio Economic Panel (SOEP, v28, 1984-2011)*. Berlin: DIW/SOEP, pp. 131–147.
- Gruber, J. (2004). Is making divorce easier bad for children? The long-run implications of unilateral divorce. *Journal of Labor Economics*, 22, 799–833.
- Härkönen, J. (2014). Birth order and educational attainment in West Germany. *European Sociological Review*, 30, 166–179.

- Hillmert, S. and Jacob, M. (2010). Selections and social selectivity on the academic track: A life course analysis of educational attainment in Germany. *Research in Social Stratification and Mobility*, 28, 59–76.
- Jonsson, J. and Gähler, M. (1997). Family dissolution, family reconstitution and children's educational careers: Recent evidence for Sweden. *Demography*, 34, 277–293.
- Kesler, C. (2003). Educational stratification and social change: Evidence from German unification. *European Sociological Review*, 19, 467–482.
- Lareau, A. (2011). *Unequal Childhoods: Class, Race, and Family Life. Second Edition*. Berkeley, CA: University of California Press.
- Mandemakers, J. and Kalmijn, M. (2014.) Do mother's and father's education condition the impact of parental divorce on child well-being? *Social Science Research*, 44, 187–199.
- McLanahan, S., Tach, L., and Schneider, D. (2013). The causal effects of father absence. *Annual Review of Sociology*, 39, 399–427.
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26, 67–82.
- Ní Bhrolcháin, M. (2001). 'Divorce effects' and causality in the social sciences. *European Sociological Review*, 17, 33–57.
- Sandefur, G. and Wells, T. (1999). Does family structure really influence educational attainment? *Social Science Research*, 28, 331–357.
- Sigle-Rushton, W., Lyngstad, T. H., Andersen, P. L., and Kravdal, Ø. (2014). Proceed with caution? Parents' union dissolution and children's educational attainment. *Journal of Marriage and Family*, 76, 161–174.
- TNS Infratest Sozialforschung (2012). SOEP 2011 – Methodenbericht zum Befragungsjahr 2011 (Welle 28) des Sozio-oekonomischen Panels. *SOEP Survey Papers* 108. Series B. Berlin: DIW/SOEP.
- Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40, 1008–1018.
- Torssander, J. and Erikson, R. (2010). Stratification and mortality: A comparison of education, class, status, and income. *European Sociological Review*, 26, 465–474.
- von Hippel, P. (2007). Regression with missing Y's: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37, 83–117.

- Wagner, G., Frick, J., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. *Schmollers Jahrbuch (Journal of Applied Social Science Studies)*, 127, 139–169.
- Westphal, S., Poortman, A.-R., and van der Lippe, T. (2014). Non-resident father-child contact across divorce cohorts: The role of father involvement during marriage. *European Sociological Review*, 30, 444–456.

## CHAPTER 4: THE CAUSAL EFFECT OF PARENTAL SEPARATION ON CHILD EDUCATION: A NEW INSTRUMENTAL VARIABLE APPROACH<sup>‡</sup>

### Introduction

Numerous studies have documented that children, whose parents separated during their childhood, have more behavioral problems, lower psychological well-being, and poorer educational outcomes than children whose families remained intact (Amato 2000, 2010; McLanahan and Sandefur 1994; McLanahan et al. 2013). These findings have been reproduced for many countries and the associations do not appear to have changed over time, even as the acceptance and occurrence of parental separation has increased (Amato and James 2010; Gähler and Palmtag 2015; Sigle-Rushton et al. 2005).

A central question in this literature, as well as in the surrounding public debate, is whether these associations reflect negative causal effects of parental separation on child well-being (Amato 2010; McLanahan et al. 2013; Ribar 2004). Researchers have used different techniques to obtain causal estimates of these effects. Each research design employed for this purpose rests on different assumptions under which the estimates can be interpreted as causal effects. As importantly, different approaches also estimate effects on different populations as a consequence of employing different research designs.

Overviews of the literature have often concluded—based on the wide heterogeneity of findings from strong effects to no effects at all—that parental separation probably has a causal negative effect on child outcomes, but that this effect is weaker than what cross-sectional estimates suggest (Amato 2010; McLanahan et al. 2013). However, it is not clear whether this conclusion is justified or whether it is merely a compromise position taken by researchers faced with diverging results. Given the interest in the topic by researchers, policy makers, and the broader public alike, further attempts to estimate the effect of parental separation on child education are warranted.

In this chapter, we propose a new instrumental variable (IV) approach to estimate the effect of parental separation on pupil's school grades at the end of compulsory school in Sweden. Theoretically, instruments are probably the most convincing method in a social scientist's toolbox to identify causal effects because they isolate that part of the variation in an explanatory variable which is truly exogenous (Angrist and Pischke 2009; Morgan and Winship 2015). We introduce a new IV to instrument parental separation: the ratio of opposite

---

<sup>‡</sup> This chapter was co-authored with Juho Härkönen (Stockholm University).

sex colleagues at the maternal workplace. Workplace sex ratios have been found to predict parental divorce and separation in Denmark (Svarer 2007) and Sweden (Åberg 2003, 2009). Furthermore, McKinnish (2004, 2007) demonstrates that sex ratios within occupations and industries affect parental divorce. We argue that it is very unlikely that sex ratios at the maternal workplace have a direct effect on child education, that is an effect which does not run through parental separation.

We analyze data on Sweden which is an interesting country when analyzing the effect of parental separation on child education because of its liberalism regarding family decisions and the generous welfare state. Previous findings indicate parental separation penalties in educational performance in Sweden, thereby confirming findings in other countries (Jonsson and Gähler 1997), although an influential study by Björklund and Sundström (2006) argues that this result is due to selection effects which disappear once family-fixed effects models are applied. Sweden also has population register data which offer the unique opportunity for implementing our IV approach.

This chapter proceeds as follows. Next, we review conceptual frameworks of the effects of parental separation on children's outcomes and discuss different methodological approaches to estimate these effects. Following to that, we introduce our estimation strategy and the assumptions we make in order to identify causal effects. Then, we describe the data and variables we employ and report our results. In the conclusion we discuss how our results can be interpreted with respect to findings presented in previous research.

## **Background**

### Assumptions Underlying the Identification of Parental Separation Effects

Scholarship on the consequences of parental separation increasingly recognizes that parental separation is both an event and a process (Amato 2000, 2010; Kim 2011). Moreover, parental separations and the circumstances surrounding them are very heterogeneous (Amato 2000, 2010). Parental separation as an event characterizes the physical separation (which does not necessarily perfectly overlap with legal divorce) of the parents and the end of the joint family. This means that both parents are no longer physically present in children's day-to-day lives, which itself means many changes to children's lives. Children can experience parental separation as confusing and feel sad or otherwise suffer because of the disruption of their

family. At the same time, some children may feel relief if the separation ends a disruptive family environment.

The separated parents can be less able to exert control over their children, which can lead to uncoordinated parenting (Amato 1993). Because the parents, too, need to adjust to the new situation and conduct their own emotional work, they can be additionally challenged in their capacity to engage in efficient parenting. Separations also have effects on parents' and their children's socio-economic circumstances. In many countries, separation is a leading predictor of poverty (Andreß and Hummelsheim 2009) and can more generally lead to a loss in parental human capital and to downward social mobility (Jonsson and Gähler 1997). This adds to the challenges faced by the separating parents and their children.

Often the process leading to parental separation begins well before the parents actually separate (Amato 2000). The dissolution of the parental union can be triggered by a specific event or it can follow from a gradual estrangement of one or both partners. The further process of dissolution is often accompanied by deteriorations in the family environment with conflicts of increasing intensity. Accordingly, the earliest effects of family dissolution on children's well-being are often observed already before the actual physical separation takes place (Kim 2011; Morrison and Cherlin 1995; Sun 2001; Sun and Li 2002). When deciding upon whether to separate or not, most parents take into consideration their children's situation and whether separating or staying in the partnership will be the best solution (Aughinbaugh et al. 2005). These and other considerations affect not only whether the parents eventually separate but also at which stage of the dissolution process they do.

Not all separations follow this pattern of deterioration in the family environment until the actual separation takes place. Many couples end partnerships characterized by at least moderate degrees of satisfaction and little conflict (Amato and Hohmann-Marriott 2007). At least judging from results of research on the reasons for divorce, the share of divorces ending less conflicting marriages has increased (De Graaf and Kalmijn 2006). Other couples continue living in unhappy—and at times conflict-ridden—partnerships for a long time, some even without ever separating (Hawkins and Booth 2005). Even if high conflict did not characterize the whole course of the partnership, the family environment may nevertheless have been poor during most of the child's life.

This underlines the heterogeneous nature of separations. The heterogeneity of separations translates into heterogeneity in the effects of parental separations. On the one hand, children, whose parents end a highly distressed and problem-ridden partnership, may benefit from the separation compared to remaining in such a problematic family environment

(Amato et al. 1995; Dronkers 1999; Hanson 1999; Jekielek 1998). On the other hand, the (physical) separation of parents can have the most deleterious effects on children's well-being if the parental partnership was characterized by low levels of conflict prior to the separation (Booth and Amato 2001; Dronkers 1999). In a recent paper, Amato and Anthony (2014) found major variation in the effects of parental separation, which were generally more negative among children with the parents with the highest propensity to divorce.

**Figure 4.1** Hypothetical scenarios of family life courses and children's educational performance around parental separation

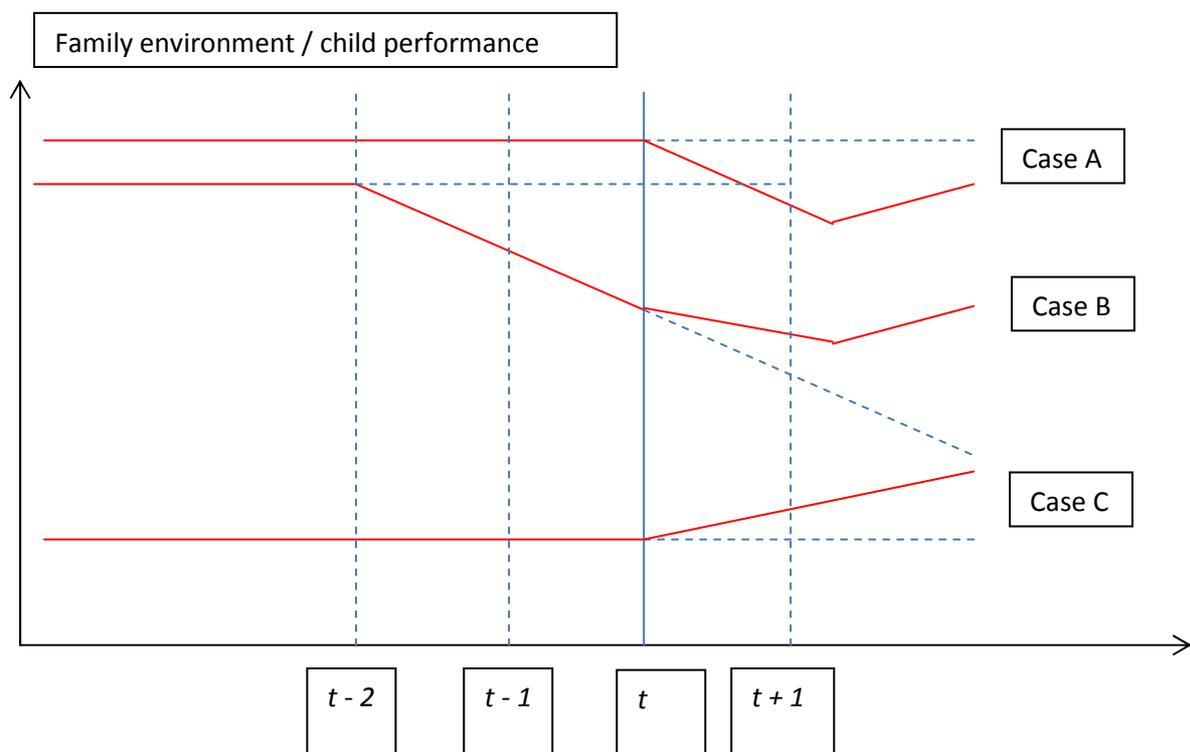


Figure 4.1 depicts the course of parental separation for children in three hypothetical situations. The y-axis represents both the quality of the family environment and the child's educational performance (assuming, for the purposes of this example, a one-to-one relationship between the two), while the x-axis represents time. The separation occurs at the solid vertical line. The first case (Case A) represents a scenario in which parental separation ends a well-functioning family; although at least one of the parents may have contemplated the separation for quite some time, this has not reflected upon the family environment nor the child's functioning. The second case (Case B) exemplifies a case in which parental separation is the culmination of deterioration in the family environment, whereas in the third case (Case C), the family has been dysfunctional for a long time, but nevertheless hitherto remained

intact. The solid lines continuing from after the separation show how the quality of the family environment, and the child's functioning, developed post-separation. In the first case, initial deterioration was followed by partial recovery. In the second case, separation ended the downward slope, and was later followed by improvement. Finally, in the third case, the separation meant an escape from the dysfunctional family. The dashed lines continuing after the separation depict counterfactual scenarios in the absence of the parental separation: in the first and third case, the positive and the adverse family environments would have continued, whereas in the second case, the family environment would have continued to deteriorate further. Despite being hypothetical scenarios, they can be seen to have their real-life equivalents (Amato 2000, 2010; Kim 2011; Sigle-Rushton et al. 2014).

These features have implications for the empirical analysis of the effects of parental separation on children's outcomes. Researchers have paid increasing attention to questions of the importance of unobserved variables for estimating causal effects of parental separation (e.g. Amato 2010; McLanahan et al. 2013; Ribar 2004) but have paid less attention to questions of which effects are estimated within each approach (Härkönen 2014; Sigle-Rushton et al. 2014). In the latter case this ultimately boils down to the question of the counterfactual reference group in an estimation strategy.

The simplest and most commonly-used strategy for studying parental separation effects on children's outcomes is to estimate effects of parental separation on children's outcomes in a cross-sectional regression framework. In such a framework—which generally relies on comparing children from separated with children from intact families—the possibilities for making causal statements depend on access to relevant confounding covariates, a requirement most often not fulfilled (McLanahan et al. 2013; Ribar 2004). The same reliance on observed control variables characterizes propensity score matching techniques and growth curve modeling.

Critics of this approach repeatedly point to parental conflict as a variable which needs to be controlled for in order to approach a causal interpretation of the parental separation estimate. Its inclusion in the model often leads to a substantial reduction in the size of the parental separation effect (e.g. Gähler and Garriga 2013; Hanson 1999). This is often interpreted as showing that parental conflict rather than the event of parental separation is responsible for children's lower educational outcomes. If the relevant control variables are included and parental conflict, or some other measure of the family environment, is measured immediately before the actual separation took place ( $t - 1$  in figure 1), the estimate from the regression model would identify the average effect of the event of separation correctly if it is

reasonable to expect that the situation at the measurement point ( $t + 1$ ) would resemble the situation at  $t - 1$  in the absence of separation. Case A and Case C would fulfill this criterion, whereas Case B would not, as in that case the family environment would have deteriorated further in the absence of the separation. Without measures to capture the deterioration of the family environment, which characterizes many dissolution processes, one would, therefore, not be able to estimate the correct effect size in this case.

If one regards parental conflict as an endogenous part of the separation process, controlling for parental conflict—at least if measured during the dissolution process—would not be a good strategy.<sup>1</sup> In such a situation, one would instead attempt to control for covariates which triggered the parental separation process and the conflicts which were part of it, for example, at  $t - 2$  in Case B in Figure 4.1. This is of course easier said than done, not least because of the lack of conceptual clarity on how to define the initiation of the partnership dissolution process. In such an unlikely scenario, one would correctly identify the average effect of the parental separation process.

As is well-known and mentioned above, a central limitation of cross-sectional regression models is that effects are biased if not all factors which lead to parental separation are observed and controlled for. Fixed effects models offer a solution to this problem (Amato and Anthony 2014; Aughinbaugh et al. 2015; Björklund and Sundström 2006; Ermisch and Francesconi 2001; Lee and McLanahan 2015; Sandefur and Wells 1999; Sigle-Rushton et al. 2014). These models are based on comparing individuals at different time points (individual-fixed effects) or siblings from the same family (family-fixed effects or sibling-difference models). Fixed effects models control for all observed and unobserved factors that remain stable across time or between siblings, respectively. These methods have become popular among researchers (for instance in the previous chapter). The identification breaks down, however, to the extent that time-varying variables or factors that vary between siblings, such as the specific traits and developments of the children, affect the risk of separation (Ermisch et al. 2004). Additional control variables have often been introduced to control for these potential sources of bias.

In the case of individual-fixed effects models, the effect of parental separation is identified if it is reasonable to expect that, in the absence of separation, the child would have experienced similar outcomes after separation (Amato and Anthony 2014; Aughinbaugh et al. 2005; Lee and McLanahan 2015). Similarly, family-fixed effects models estimate the effect if it is reasonable to expect that the older sibling experienced conditions prior to the parental separation which the younger sibling would have experienced in its absence (see previous

chapter and Sigle-Rushton et al. 2014). In both cases, the assumption is more likely to be met if the family environment was stable during the years leading to the separation (Case A and Case C in Figure 1) but less likely if the separation was a culmination of a relatively brief process of deterioration in the family environment (cf. Sigle-Rushton et al. 2014).

These assumptions have been relaxed in individual-fixed effects models by taking a triple-differencing approach (Sanz-de-Galdeano and Vuri 2007)<sup>2</sup> or by tracking the development in the outcome variable before and after the separation among children who experience the event (Aughinbaugh et al. 2005). Essentially, however, fixed effects approaches estimate the effect of the event of separation, although by tracking the development in the outcome variable pre- and post-separation can inform on the separation process. Because the counterfactual in fixed effects models is based on assuming that pre-separation child outcomes ( $t - n$ ) are informative of the post-separation ones ( $t + n$ ) (individual-fixed effects) or the outcomes of the older sibling are informative of the counterfactual outcomes of the younger one at a similar age (family-fixed effects), inherently time-dependent fixed effects estimates are probably best interpreted as telling about whether postponement of the physical separation would have been beneficial to the child or not (Sigle-Rushton et al. 2014).

The purpose of the discussion above was to highlight how different approaches to estimate the effects of parental separation on educational outcomes can lead to insights on different aspects of parental separations, which are often not one-shot events and show a great deal of heterogeneity. Much of the methodological discussion on estimating these effects has focused primarily on selection into parental separation, in other words, on the possibilities to control for omitted variable bias, whereas the question for which population effects are estimated has received less attention. We, however, argue that both questions are of similar importance and intrinsically linked to the choice of research design.

Below we outline our research design. In describing it we not only discuss the assumptions we need to make in order to give a causal interpretation to the estimates obtained using this method but also try to clarify who belongs to the population for whom we estimate a causal effect.

## Heterogeneity in the Effects of Parental Separation on Child Education

Research on the consequences of parental separation has been criticized for not taking into account that the effects of parental separation may be unequally distributed within the

population of children affected by parental separation (Amato 2010; Amato and Anthony 2014; Lee and McLanahan 2015). To some degree, the IV approach already includes a perspective that the effects of parental separation may vary. The effect is identified in the population which is affected by the instrument (in our case those relationships affected by the ratio of opposite sex colleagues at the maternal workplace). This perspective does not, however, explain by which characteristic the effects of parental separation vary. To shed light on this question, we analyze the heterogeneity of effects along two sociologically important dimensions: child gender and social origin.

First, heterogeneity in separation effects can occur with respect to child gender. Until recently the consensus, based on empirical research testing this prediction largely using cross-sectional research designs, seemed to be that parental separation effects do not differ for girls and boys (Amato 2010). Contrary to that, Lee and McLanahan (2015), using data on the United States and individual-fixed effects models, find that the cognitive performance of girls is more negatively affected by parental separation than the performance of boys. We are interested in finding out whether we can reproduce this finding using our IV approach.

Second, differences in the association between parental separation and child education between social origin groups have been recently investigated in several studies. These studies, largely not employing any causal identification strategy in order to control for selection into parental separation, lead to mixed results with most studies finding the negative effects of parental separation being more pronounced in families with a high than in families with a low level of parental education (Albertini and Dronkers 2009; Augustine 2014; Mandemakers and Kalmijn 2014), although other studies report the opposite result (Bernardi and Radl 2014; Bernardi et al. 2014). In Chapter 2, I employ a research design which controls for unobserved variables (family-fixed effects models) and find that the effects of parental separation on educational outcomes are concentrated in families with a low level of parental education.

## **Research Design**

We use exogenous variation on the sex composition at the maternal workplace at the time of the birth of the child to identify the effect of parental separation in an instrumental variable (IV) framework. In an IV setting, an exogenous source of variation, which is not directly related to the outcome of interest (here, children's school grades) induces some parents to separate but leaves other parental relationships intact. The challenge is to find an instrument which affects parental separation but has an effect on children's outcomes only through its

effect on parental separation. Below, we discuss the assumptions underlying IV estimation as well as the effect which our approach estimates and the population to which it generalizes to.

IV estimation requires an instrument which affects parental separation but has an effect on children's outcomes only through its effect on parental separation. For this purpose, we use information on the ratio of co-workers of the opposite sex at the maternal workplace. Sex ratios at the workplace have been identified as a predictor of parental divorce and separation in Denmark (Svarer 2007) and in Sweden (Åberg 2003, 2009). Similarly, McKinnish (2004, 2007) shows that sex ratios within industries and occupations predict divorce. These studies have led to the conclusion that workplace sex ratios do, indeed, have an effect on divorce and separation independent of various control variables. A workplace with a higher ratio of co-workers of the opposite sex provides parents with opportunities and possibly, temptations to engage in romantic encounters with one's colleagues, increasing the risk of separation. Furthermore, workplaces constitute a marriage market segment also to those already in a relationship. Svarer (2007) finds that workplace sex ratios do not affect entry into marriage for singles, implying that this marriage market segment is particularly strong for those already in a relationship, possibly due to higher search costs for (alternative) partners. We focus on the maternal workplace since our explorative analysis found only weak effects of the ratio of opposite sex colleagues at the paternal workplace on parental separation. Since a strong instrument is required to obtain unbiased second stage estimates (that is, estimates of the effects of parental separation on school grades), we dropped the ratio of opposite sex colleagues at the paternal workplace as a possible instrument. We return to the issue of weak instruments below when discussing our results.

Besides affecting the risk of separation, the central identifying assumption allowing us to interpret the estimates of our IV models as causal effects is that the maternal workplace sex ratio is related to children's educational success only through its effects on the parental propensity to separate but does not affect child outcomes through other pathways. One way of minimizing the risk of possible violations of this assumption is to include control variables which are related both to the sorting of mothers into sex or gender segregated workplaces as well as to educational outcomes. We include a number of such control variables, including the mother's educational attainment, her field of study, and her region.

It is still possible that residual confounding remains, either because of unobserved factors related both to sorting into workplaces and educational outcomes, or because maternal workplace sex ratios have a direct effect on school outcomes independently of parental separation. The exclusion restriction assumption—namely that the instrument affects the

outcome only through the treatment—is inherently untestable (Imbens and Angrist 1994). We argue that in our study the probability for the latter source of confounding is small. However, one such possibility is that a higher ratio of men at the mother’s workplace may increase her risk of a romantic affair with a co-worker without always leading to a separation. In this case, the family environment can deteriorate but the family remains intact. Through this pathway, maternal workplace sex ratios can affect children’s outcomes even in families that remained intact. Given that we measure the mother’s workplace sex ratio at the time of birth, one can argue that most parents whose long-term relationship quality was affected by infidelity would eventually have separated by the time the grades were measured, even if the partnership survived the infidelity in the short-run. If the shock to partnership was short term, the fifteen years between the measurement of the maternal workplace sex ratio and grades allows for time to recover from any negative effects. In any case, the possibility of remaining bias—either through unobserved factors affecting the sorting into occupations or through a direct effect of workplace sex ratios on school grades—can never be completely ruled out. This bias will be stronger if the instrument is weak, even in large data sets such as the one we employ in this analysis (Small and Rosenbaum 2008).

Even when all the identifying assumptions are met, IV estimation identifies the effects of parental separation only for a specific sub-population. Under the LATE interpretation of IV estimates, our estimates represent the effect of parental separation on the children whose parents separated because of the specific sex ratios present at their mothers’ workplaces but who would not have separated otherwise (Imbens 2010; Imbens and Angrist 1994). This effect does not, thus, necessarily generalize to those whose parents separated for other reasons. However, given the recent emphasis on heterogeneous effects of parental separation, our estimates will be informative.

To better understand to which sub-population the effects can generalize to, it is important to understand who the parents are whose partnership dynamics will be affected by the workplace sex ratio (the “compliers” in IV parlance) and what kinds of separations may be involved. First, we distinguish them from the “always-takers” and the “never-takers”, that is, those who would separate in any case or those who would never separate. The first group is likely to be over-represented by parents, whose partnerships are characterized by such toxicity that separation is next to inevitable. In Figure 4.1, these would be characterized especially by Case C, and to a lesser extent by Case B. Another possible, yet probably smaller, sub-group of parents in this group are those with extremely low relationship commitment. The “never-takers”, on the other hand, would chiefly consist of high-quality partnerships and, those with

very high moral or other barriers to separation, although this will be a smaller group in a country like Sweden, which has liberal attitudes to parental separations (Rijken and Liefbroer 2012).<sup>3</sup>

Based on previous findings on workplace sex ratios and separation, we argue that our group of compliers consists importantly of parents who find a new partner from their workplace, and who would not have found a new partner (or other reason to separate) had they worked in a workplace with a different ratio of opposite sex colleagues. Given Svarer's (2007) argument that workplaces are particularly important partner markets for those already in a partnership, this interpretation is feasible. It is likely that many of those who end up separating because they find a new partner at the workplace lived in reasonably well-functioning partnerships because they would not have separated otherwise. This does not need to mean that the parents were completely happy with their partnership; most would otherwise not have engaged in an affair to begin with. However, the condition that the parents would have stayed together were it not for the workplace sex ratio does rule out separations in which getting involved with a co-worker was a final means to get out from a partnership which one was already determined to end. This adds to the feasibility of our interpretation.

There are reasons to believe that the experiences of children whose parents separated because they found a new partner from the workplace are more negative than on average. First, infidelity often leads to severe family conflict and falls in parental well-being, which can make separations involving new partners especially difficult. Cano and O'Leary (2000), for example, reported markedly increased incidences of depression among women whose (former) husbands cheated on them. These effects can linger long after the separation.

Second, as mentioned above, many separations caused by finding a new partner from the workplace probably dissolve previously relatively well-functioning relationships that would have remained intact were it not for the biased workplace sex ratios. Third-party involvement can lead to a rather rapid dissolution of a family, either because the partner who became romantically involved or the other partner (once he or she becomes aware of the affair) wants to end the relationship quickly. This could correspond to a separation exemplified by Case A in Figure 1 (even if the family environment at the baseline was lower than in Case A). If unexpected separations have more severe effects on children than those ending a long-standing conflicting family environment—because in the former case the separation leads to deterioration of the family environment, whereas in the latter this is not necessarily the case—these types of separations can have more severe effects on the children. In other cases, even if workplace infidelity does not lead to an instant separation, it can lead to

a deterioration of the parental partnership (and the children's outcomes), which eventually leads to separation. Case B would be an example. In both these cases, we argue that we identify the "total" effect of parental separation—including the potential pre-separation conflicts and deterioration of the home environment—as the workplace partner market is in both cases the triggering event which eventually leads to the separation.

Third, the effects of parental separation for the complier group can be expected to be additionally severe because of the post-separation family arrangements. These can include the new partner, whose presence can have negative effects (Hanson et al. 1996), and continuing conflicts between the biological parents of the child.

Summing up, we argue that for the children of compliers, our IV strategy identifies the total effect of experiencing parental separation and the time following separation. By total we mean the effect of both the process and the event of the separation but also the effect of a possible recovery phase since we measure educational performance at a later point in time. We argue that our approach also identifies the effect of the process of separation because our IV triggers the separation process, which would not have occurred otherwise. Separations triggered by biased sex ratios are likely to dissolve relatively well-functioning families which would not have separated otherwise and separations involving third parties are likely to involve conflictual separation processes and post-separation family environments which may have adverse effects.

## **Data and Variables**

### Data and Sample Selection

We analyze data on seven birth cohorts of children born in Sweden between 1990 and 1996 on whom we have information on their grade point averages (GPA) in ninth grade in the years 2006 to 2012. The data come from the Sweden in Time - Activities and Relations (STAR) register database which is maintained at Stockholm University.

Our IV uses information on the maternal workplace. The workplace is the unit where people work within private and public organizations. This information is not available for all workplaces but for all public workplaces, all bigger private workplaces, and a random sample of smaller private workplaces. Because of this data limitation, we cannot estimate the effects of the workplace instrument for our entire school cohort but only for a very large subsample.

However, because of the random nature of the sampling of the workplaces this should not bias our estimates.

Since we are interested in estimating the effects of parental separation, and not the effects of single parenthood, we restrict the sample to children born into two-parent families. This means that we drop children who were born into single-parent families from the analysis sample. In addition, we only include those children with parents working at a workplace with 5 and more employees in the analysis sample. At smaller workplaces, the sex ratio at the workplace is a worse predictor of parental separation because small changes in the number of persons at the workplace can drastically change the sex ratio (Svarer 2007).

A limitation of our approach is that we can only estimate the effects of parental separation in a sample of children with working mothers. We do not have information on the workplace for mothers who are unemployed or inactive. Due to this sample selection, we may possibly underestimate the effects of parental separation. However, we expect that such bias will be rather small due to the, in particular in comparison to other countries, rather small number of unemployed and inactive mothers in Sweden.

## Variables

We measure parental separation based on the household structure in which the children lived during their childhood (Thomson and Eriksson 2013). We define parental separation as has happened when a child at any year prior to receiving his GPA did not live in the same dwelling with both of her or his parents.

We measure educational performance through the Grade Point Average (GPA, “*meritvärde*”) in the ninth grade of compulsory school. This is a continuous variable and commonly used in research on educational inequalities in Sweden (Erikson and Rudolphi 2010; Rudolphi 2013). According to Rudolphi (2013) more than 75% of inequality in educational opportunity is covered by GPA scores so that final educational attainment highly correlates with GPA. Children are around 16 years old during the time GPA is measured. We standardize GPA so that it has a mean of 0 and a standard deviation of 1. This allows us to interpret the effects of parental separation on child GPA in terms of standard deviations, making it easier to compare our results to findings from research on other countries.

Our instrumental variable (IV) is a measure of the ratio of opposite sex colleagues at the maternal workplace. In order to construct this variable, we rely on information on the workplaces of the parents. A workplace is the physical location where people work. Using the

ID of workplaces, we can calculate workplace sex ratios and then construct a measure of the sex ratio of opposite sex colleagues. In order to avoid endogeneity due to reverse causality, we use information on the maternal workplace in the child's birth year. Since we focus on parental separations after birth, we can ensure through this definition that the information on the sex composition of the workplace precedes parental separation.

As control variables we use gender and birth order in all models. Even though these variables are only weakly related to our instrument, they help to explain the variance in GPA and to reduce the standard errors of our IV estimates. Furthermore, the influence of these variables on GPA provides us with estimates of the effects of other important factors which influence child education to which we can compare the size of the parental separation penalty.

More importantly, in order to ensure the exogeneity of our IV, we control for maternal education, field of study, workplace size, and the region in which the mother lives. These controls are supposed to take up the effects of sex segregation into occupations and any selection into workplaces with certain sex ratios. In addition, we estimate effects within regions in order to control for regional variations in relationship markets. Because of these controls we essentially estimate within-education-groups, within-educational-fields, and within-region estimates. For that reason, our IV strategy only relies on the assumption that the distribution to workplaces within levels of education, educational fields, and regions is random. We are not aware of any evidence speaking against this assumption.

We use information on parental education in order to analyze whether the effects differ between children from high and low social origin families. We define families to have a high level of parental education if at least one of the parents attended the longer track in secondary school (*Gymnasium*).

**Table 4.1** Descriptive statistics for variables in the sample used in the analysis

<b>Variable</b>	<b>Mean</b>	<b>Std. dev.</b>	<b>Min.</b>	<b>Max.</b>	<b>N</b>
Grade Point Average (GPA)	0.20	0.87	-3.12	1.74	348,034
Female	0.49	0.50	0	1	348,034
Parental separation	0.31	0.46	0	1	348,034
Birth order	1.84	0.90	1	16	348,034
Ratio opposite sex co-workers at maternal workplace	0.32	0.25	0	0.99	348,034
Size maternal workplace	687.06	1557.43	5	9697	348,034
High parental education	0.59	0.49	0	1	348,034
Ratio opposite sex co-workers in maternal industry branch	0.37	0.23	0.06	0.94	347,646

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

Table 4.1 reports descriptive statistics on the school cohorts which we use for our analysis. About 31% of the children in our sample have experienced parental separation before GPA is measured. This number is in line with previous research estimating the frequency of the experience of parental separation for these birth cohorts (Thomson and Eriksson 2013).

## Results

### Baseline Estimates: OLS Regressions on GPA at Age 16

We start the analysis by presenting results on the association between parental separation and GPA in ninth grade. Table 4.2 reports OLS regression models with GPA as the outcome variable and parental separation as our explanatory variable. These results serve as the baseline estimates to which we compare our IV estimates. They also make our analysis comparable to earlier research which most of the time has reported estimates of the association between parental separation and educational outcomes of a similar size.

**Table 4.2** OLS regression models on the association between parental separation and child GPA

	Grade Point Average (GPA)					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Parental separation	-0.36** (0.00)	-0.27** (0.00)	-0.27** (0.00)	-0.28** (0.00)	-0.29** (0.00)	-0.24** (0.00)
Female	0.33** (0.00)	0.34** (0.00)			0.34** (0.00)	0.34** (0.00)
Birth order	-0.12** (0.00)	-0.10** (0.00)	-0.10** (0.00)	-0.11** (0.00)	-0.11** (0.00)	-0.10** (0.00)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	348,034	348,034	177,831	170,203	142,776	205,258

*Note:* OLS regression models. Standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

Model 1 in Table 4.2 reports the association between parental separation and GPA, only controlling for child sex, cohort, and birth order, as about one third of a standard deviation.

Conditioning on maternal level of education, field of study, size of the maternal workplace, and region reduces slightly the size of the parental separation penalty. According to Model 2, children who experienced a parental separation score about a quarter of a standard deviation lower in their GPA than children whose parents stayed together. This result suggests a large, negative effect of family structure on child education. The negative effect of having experienced a parental separation is of a similar size as the negative effects of having two to three older siblings or being male.

Model 3 to 6 analyze the heterogeneity in the association between parental separation and GPA. Overall, we find no evidence that there is heterogeneity by child gender or parental education in the association between these two variables. The coefficients in the male and female subsample are identical. The negative association is slightly stronger in the sample of children with low educated parents than in the sample of children with highly educated parents. However, the size of the difference is very small.

#### First Stage Results of the Instrumental Variable Regression

In order to be able to use the ratio of opposite sex colleagues at the maternal workplace as an IV for parental separation, we have to demonstrate that it affects the probability of parents to separate. In Table 4.3 we report Linear Probability Models which predict whether a child in our sample has experienced parental separation. We prefer to report LPM because of the easier interpretation of these models and because they can be integrated easily into a 2SLS framework (Angrist and Pischke 2009).

We find a positive and statistically significant effect of the ratio of male colleagues at the mother's workplace on separation risk. We find a clear difference in the strength of the effect in low educated compared to highly educated households: the effect is almost three times stronger among the less educated. This socio-economic difference in the strength of the first stage may be due to that low educated partners separate more often than highly educated parents. Recent findings indicate that this negative educational gradient of divorce is due to higher barriers to divorce among the highly educated leading to a higher threshold for divorcing (Boertien and Härkönen 2014). These barriers can prevent that marriage markets for alternative partners among the highly educated lead to a family dissolution, whereas the less educated who have fewer barriers are more susceptible to their influence. In IV parlance, the share of compliers is larger among the low than among the highly educated parents (Morgan and Winship 2015: 321). The Cragg-Donald Wald F-statistic, a test of the strength of the

instrument, is in all regressions larger than 10 (Stock and Staiger 1997) and or even 16.38, which can be considered a more conservative test (Stock and Yogo 2005). Hence, the usual test statistics used for IV regressions suggest that our instruments are valid. Nevertheless, the effect size among the highly educated parents suggests that the instrument is weak in this group. For the children with highly educated parents, going from an all-female to an almost all-male workplace increases the probability of parental separation by less than three percentage points. Given our large sample size, this estimate is statistically significant but may not be strong enough to prevent problems with weak instruments (Bound et al. 1995; Small and Rosenbaum 2008). We return to these considerations in the interpretation of the second stage results and in the discussion.

**Table 4.3** Linear Probability Models of the effects of workplace sex ratio on parental separation

	<b>Parental separation</b>					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Ratio opposite sex colleagues at maternal workplace	0.069** (0.003)	0.048** (0.004)	0.047** (0.005)	0.049** (0.005)	0.072** (0.005)	0.028** (0.005)
Female	0.008** (0.002)	0.007** (0.002)			0.010** (0.003)	0.005* (0.002)
Birth order	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	-0.000 (0.001)	-0.005** (0.001)	0.004** (0.001)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	348,034	348,034	177,831	170,203	142,776	205,258
<i>Cragg-Donald</i>	475.452	192.458	94.834	97.908	183.987	36.716
<i>Wald F statistic</i>						

*Note:* Linear Probability Models. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

#### Instrumental Variable Estimates of the Effects of Parental Separation on Child Education

Table 4.4 reports the effects of parental separation on educational outcomes using the ratio of opposite sex co-workers at the maternal workplace to instrument parental separation. The estimates for the full sample report no negative effect of parental separation on GPA.

Contrary to that, the estimate is positive suggesting that, if at all, children who experience parental separation are rather positively influenced by their parent's separation (although these positive effects do disappear in the first robustness check reported below). A possible explanation is that, unlike expected, the separations induced by workplace sex ratios end inherently dysfunctional families and mean that the children can leave a poisonous family environment. One issue may also be that the timing of the separation plays a role. Since we measure sex ratio at child's birth and GPA around age 16 there can be a lot of time which has passed between parental separation and the measurement of educational performance. This time may have led to a recovery of the educational performance of most children.

**Table 4.4** 2SLS estimates of the effects of parental separation on child GPA, using the ratio of opposite sex colleagues at the maternal workplace as an IV for parental separation

	Grade Point Average (GPA)					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Parental separation	0.76** (0.10)	0.47** (0.14)	0.28 (0.19)	0.66** (0.20)	-0.32* (0.13)	2.25** (0.51)
Female	0.32** (0.00)	0.33** (0.00)			0.34** (0.00)	0.33** (0.01)
Birth order	-0.12** (0.00)	-0.10** (0.00)	-0.10** (0.00)	-0.11** (0.00)	-0.11** (0.00)	-0.11** (0.00)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	348,034	348,034	177,831	170,203	142,776	205,258

*Note:* Second stages of 2SLS IV regression models. Parental separation instrumented by ratio of opposite sex co-workers at the maternal workplace. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

Model 2 and 3 in Table 4.4 report results separately by child gender. These estimates indicate more positive effects of parental separation on GPA for girls than for boys. However, the difference is small. In any case, for both groups, the estimates are actually positive and go, hence, in the same direction. Because neither for girls nor for boys the effect becomes negative, we conclude that the parental separation penalty does not vary by child gender.

Clearly different are, however, the results if we split the sample by parental education. Models 4 and 5 in Table 4.4 report results separately for children from highly and low educated parents. Our results indicate that there is one group of children which is actually

negatively affected by parental separation: children from families with a low level of parental education. For this group we find a strong negative effect of parental separation of about one third of a standard deviation. This estimate is close to the estimate for this group in the baseline model which does not control for selection into parental separation on unobserved variables. Contrary to that, we find no evidence of a negative effect of parental separation on GPA for children from highly educated families.

These results suggest that there is important socio-economic heterogeneity in parental separation effects. We show that the negative effect is more pronounced for children from families with a low level of parental education. Therefore, the average effect of parental separation misrepresents the true consequences of parental separation. This result is fully in line with the finding of a concentration of the parental separation penalty in families with a low level of parental education using family-fixed effects models and German data in Chapter 3. The overall heterogeneity in effects, including positive effects for some children, are also in line with Amato and Anthony's (2014) findings.

However, it should also be noted that one alternative interpretation of this finding is possible. The first stage results in Model 5 and 6 in Table 4.3 show that parental separation is clearly more affected by the sex ratio at the maternal workplace in families with a low than in families with a high level of parental separation. The instrument, hence, is more powerful for this first group of children than for the second. The weakness of the instrument in the highly educated group may bias the estimate, leading to the unexpected very large effect size in this group (Bound et al. 1995; Morgan and Winship 2015). A weak instrument will also inflate any bias from remaining confounding, regardless of data size, which can further contribute to the estimate size (Small and Rosenbaum 2008). Because of these considerations, we have most confidence in the IV estimate in the group of children with low educated parents, where we find a negative effect of parental separation on the GPA of a magnitude of one third of a standard deviation. Although large, this estimate is in line with the distribution of effect sizes reported by Amato and Anthony (2014). Again, it is worth keeping in mind that we estimate the effect of parental separation in population of parents whose separation was influenced by the workplace sex ratio, which can locate the effects in a likewise limited part of the distribution of effects.

Table A4.1 in the appendix reports the results of gender differences within the group of children with low educated parents. We find that a parental separation reduces the GPA of boys with low educated parents by about 0.44 standard deviations and of girls with low educated parents by about 0.21 standard deviations. This result, opposed to recent findings

reported by Lee and McLanahan (2015), provides some evidence that boys are more negatively affected by parental separation than girls. The difference is, however, small.

### Robustness Checks

In order to provide some indication of the robustness of our results, we report two robustness checks with alternative IV specifications. First, we report results using the interaction between workplace size and ratio of opposite sex colleagues as a second instrument. The rationale behind this robustness check is that the effects of the ratio of opposite sex colleagues may vary with the workplace size and including this interaction may, therefore, strengthen our first stage estimation. These first stages results are reported in Table A4.2 in the appendix.

**Table 4.5** 2SLS estimates of the effects of parental separation on child GPA, using the ratio of opposite sex colleagues at the maternal workplace and the interaction with maternal workplace size as IV for parental separation

	Grade Point Average (GPA)					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Parental separation	-0.06 (0.08)	-0.01 (0.12)	-0.23 (0.16)	0.23 (0.18)	-0.56** (0.12)	1.55** (0.40)
Female	0.33** (0.00)	0.34** (0.00)			0.34** (0.00)	0.33** (0.01)
Birth order	-0.12** (0.00)	-0.10** (0.00)	-0.10** (0.00)	-0.11** (0.00)	-0.11** (0.00)	-0.11** (0.00)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Control for maternal workplace size	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	348,034	348,034	177,831	170,203	142,776	205,258

*Note:* Second stages of 2SLS IV regression models. Parental separation instrumented by ratio of opposite sex co-workers at the maternal workplace and the interaction between this ratio and workplace size. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

The second stage results are reported in Table 4.5. These results lead to more negative parental separation effects than the main specification. In particular, these results suggest that there is no statistically significant or substantial positive effect of parental separation at the population level. The central result, however, that parental separation affects negatively the

GPA of children with low educated parents is reproduced in this robustness check and the corresponding estimate is even more negative than in the specification reported above.

Second, we use the ratio of opposite sex co-workers in specific industry branches as an alternative instrument for parental separation (McKinnish 2004, 2007). The idea is that this instrument relies on a different assumption than the instrument which we use in the main specification presented above. Parents may not be able to change the industry they work in during their professional career, making it unlikely that this variable changes if parents are not satisfied with their relationships.

**Table 4.6** 2SLS estimates of the effects of parental separation on child GPA, using the ratio of opposite sex colleagues in the maternal industry branch as an IV for parental separation

	Grade Point Average (GPA)					
	(1) Full sample	(2) Full sample	(2) Male	(3) Female	(4) Low parental education	(5) High parental education
Parental separation	-0.11 (0.08)	0.32* (0.16)	0.09 (0.21)	0.55* (0.23)	-0.46** (0.14)	2.53** (0.70)
Female	0.33** (0.00)	0.33** (0.00)			0.34** (0.00)	0.33** (0.01)
Birth order	-0.12** (0.00)	-0.10** (0.00)	-0.10** (0.00)	-0.11** (0.00)	-0.11** (0.00)	-0.11** (0.00)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	347,646	347,646	177,623	170,023	142,608	205,038

*Note:* Second stages of 2SLS IV regression models. Parental separation instrumented by ratio of opposite sex colleagues in the maternal industry branch. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

Table 4.6 reports the results of the second stage of the IV regression, whilst the first stage estimates are reported in Table A4.3 in the appendix. The results of this robustness check are surprisingly consistent with our previous results. Again, we find no general negative effect of parental separation. However, the situation is different for children from low educated parents. For them, parental separation has a considerably negative effect, in this model about two fifths of a standard deviation. For that reason, the main result of our analysis, the strong socio-economic heterogeneity in parental separation effects is supported by using this alternative IV.

## Discussion and Conclusion

There is no doubt that parental separation during childhood is associated with lower educational outcomes. However, the more important question, especially in terms of relevance for policy making, is arguably whether family structure during childhood has a causal effect on child education. In this chapter we have tested this contested hypothesis using a novel approach. The results of our IV approach suggest that there is no average negative causal effect of parental separation during childhood on education for the population which is affected by our instrument. Hence, our results using a new method question the conclusion that parental separation has, on average, a negative causal effect on child education (McLanahan et al. 2013).

However, we have also shown that this conclusion only holds at the population level but changes once the heterogeneity of separation effects is taken into account. In this respect, we found that there is a negative effect of parental separation in families with a low level of parental education. Within this subgroup, we found that boys are more negatively affected by parental separation than girls—a result which goes in the opposite direction than recent findings by Lee and McLanahan (2015) using U. S. data. There may be more heterogeneity in effects which we have not tested for. Answering these questions may be important if a policy is considered which should help children who are negatively affected by a parental separation to deal with its consequences with respect to their life chances. The identification of a causal effect puts us in a position to investigate these questions.

Nevertheless, the limitations of the approach we have used in this chapter have to be kept in mind. As we discussed, we estimate the effect of parental separation for those children whose parents have separated due to the sex composition at the maternal workplace. We believe that this group may experience a more negative effect of parental separation than children who experience parental separation for other reasons. This effect could be particularly strong for low educated parents which may be why we find a concentration of the negative effects of parental separation in that group of children. It is quite likely that for other parts of the population of children who experience parental separation, the effect on their educational outcomes is less negative. In order to identify effects for these populations, other research designs are needed. In addition, the weakness of our IV, in particular in the group of children with highly educated parents, may bias the estimates. This is why we have the most confidence in the IV estimates in the sample of children with low educated parents because

for this group maternal workplace and industry branch sex ratios were the most predictive of parental separation.

Furthermore, we have only analyzed data on seven birth cohorts in one country. Effects in other countries and for other cohorts may turn out differently which is why generalizations from this study should be done cautiously. For instance, results on the social origin heterogeneity in the association between parental separation and child education are mixed (Albertini and Dronkers 2009; Augustine 2014; Bernardi and Radl 2014; Bernardi et al. 2014; Mandemakers and Kalmijn 2014). This may be because most of these previous studies have not controlled for selection into parental separation. But the results could also diverge for other reasons. Previous studies differ in the countries and cohorts they analyze as well as in the way educational outcomes and social origin are measured. However, without acknowledging that selection into parental separation may play a role it is difficult to bring the contradictory results of previous studies in line with each other. In addition, we show differences between our OLS and IV estimates, suggesting that selection into parental separation on unobserved characteristics may indeed differ between social origin groups. This finding is also in line with the results reported in Chapter 2 which used a different research design and data on a different country but still finds a concentration of the parental separation penalty in families with a low level of parental education.

Finally, the question remains which mechanisms bring about the effect of parental separation on child education (Amato 1993; Jonsson and Gähler 1997). The identification of a causal effect as a starting point allows researchers to analyze which mechanisms bring about the causal effect and we agree that a complete understanding of the effects of parental separation requires also testing mechanisms bringing about the effect (Morgan and Winship 2015). We cannot do this with the data we used in this paper because of a lack of indicators which could be used to operationalize the underlying mechanisms. We hope that further research will take up this challenge.

## ENDNOTES

1. Additionally, one needs to be clear about when parental conflict is measured. If it refers to the immediate pre-separation period, it improves the possibility for making causal claims of the effects of the separation event. If asked retrospectively, it may refer also to the post-separation period.

2. These rely on the assumption that the unmeasured factors affecting both the separation and the outcome grow at a constant rate during the observation period.
3. The fourth group, the "defiers" would act exactly the opposite way to what we would expect, that is, be more prone to separate if working in a workplace with more same sex than opposite sex co-workers. If the partner market explanation is correct, this would probably refer to those finding a same sex partner after being in an opposite sex relationship with children. This is most likely a small group so that we can argue that there are nearly no defiers.

## REFERENCES

- Åberg, Y. (2003). *Social Interactions: Studies of Contextual Effects and Endogenous Processes*. Ph.D. dissertation. Department of Sociology, Stockholm University.
- Åberg, Y. (2009). The contagiousness of divorce. In P. Hedström and P. Bearman (Eds.), *The Oxford Handbook of Analytical Sociology* (pp. 342–364). Oxford: Oxford University Press.
- Albertini, M. and Dronkers, J. (2009). Effects of divorce on children's educational attainment in a Mediterranean and catholic society. *European Societies*, 11, 137–159.
- Amato, P. R. (1993). Children's adjustment to divorce: Theories, hypotheses, and empirical support. *Journal of Marriage and Family*, 55, 23–38.
- Amato, P. R. (2000). The consequences of divorce for adults and children. *Journal of Marriage and Family*, 62, 1269–1287.
- Amato, P. R. (2010). Research on divorce: Continuing developments and new trends. *Journal of Marriage and Family*, 72, 650–666.
- Amato, P. R. and Anthony, C. J. (2014). Estimating the effects of parental divorce and death with fixed effects models. *Journal of Marriage and Family*, 76, 370–386.
- Amato, P. R. and Hohmann-Marriott, B. (2007). A comparison of high- and low-distress marriages that end in divorce. *Journal of Marriage and Family*, 69, 621–638.
- Amato, P. R. and James, S. (2010). Divorce in Europe and the United States: Commonalities and differences across nations. *Family Science*, 1, 2–13.
- Amato, P. R., Loomis, L. S., and Booth, A. (1995). Parental divorce, marital conflict, and offspring well-being during early adulthood. *Social Forces*, 73, 895–915.
- Andreß, H.-J. and Hummelsheim, D. (Eds.) (2009). *When Marriage Ends: Economic and Social Consequences of Partnership Dissolution*. Cheltenham: Edward Elgar.

- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Aughinbaugh, A., Pierret, C., and Rothstein, D. (2005). The impact of family structure transitions on youth achievement: Evidence from the children of the NLSY79. *Demography*, 42, 447–468.
- Augustine, J. M. (2014). Maternal education and the unequal significance of family structure for children's early achievement. *Social Forces*, 93, 687–718.
- Bernardi, F. and Radl, J. (2014). The long-term consequences of parental divorce for children's educational attainment. *Demographic Research*, 30, 1653–1680.
- Bernardi, F., Boertien, D., and Popova, D. (2014). Differential effects of parental separation on child outcomes: Are children from higher social backgrounds affected more? *Working Paper European University Institute Max Weber Program*, No. 2014-06.
- Björklund, A. and Sundström, M. (2006). Parental separation and children's educational attainment: A siblings analysis on Swedish Register Data. *Economica*, 73, 605–624.
- Boertien, D. and Härkönen, J. (2014). Less education, more divorce: Explaining the inverse relationship between women's education and divorce. *Stockholm Research Reports in Demography*, No. 2014-11.
- Booth, A. and Amato, P. R. (2001). Parental predivorce relations and offspring postdivorce well-being. *Journal of Marriage and Family*, 63, 197–212.
- Bound, J., Jaeger, D. A., and Baker, R. M. (1995.) Problems with instrumental variables estimation when the correlation between the instruments and the endogenous variable is weak. *Journal of the American Statistical Association*, 90, 443–450.
- Cano, A. and O'Leary, K. D. (2000). Infidelity and separations precipitate major depressive episodes and symptoms of nonspecific depression and anxiety. *Journal of Consulting and Clinical Psychology*, 68, 774–781.
- De Graaf, P. and Kalmijn, M. (2006). Divorce motives in a period of rising divorce: Evidence from a Dutch life-history survey. *Journal of Family Issues*, 27, 483–505.
- Dronkers, J. (1999). The effects of parental conflicts and divorce on the well-being of pupils in Dutch secondary education. *European Sociological Review*, 15, 195–212.
- Erikson, R. and Rudolphi, F. (2010). Change in social selection to upper secondary school: Primary and secondary effects in Sweden. *European Sociological Review*, 26, 291–305.
- Ermisch, J. and Francesconi, M. (2001). Family structure and children's achievements. *Journal of Population Economics*, 14, 249–270.

- Ermisch, J., Francesconi, M., and Pevalin, D. J. (2004). Parental partnership and joblessness in childhood and their influence on young people's outcomes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 167, 69–101.
- Gähler, M. and Garriga, A. (2013). Has the association between parental divorce and young adults' psychological problems changed over time? Evidence from Sweden, 1968-2000. *Journal of Family Issues*, 34, 784–808.
- Gähler, M. and Palmtag, E.-L. (2015). Parental divorce, psychological well-being and educational attainment: Changed experience, unchanged effect among Swedes 1892-1991. *Social Indicators Research*, 123, 601–623.
- Hanson, T. L., McLanahan, S., and Thomson, E. (1996). Double jeopardy: Parental conflict and stepfamily outcomes for children. *Journal of Marriage and the Family*, 58, 141–154.
- Hanson, T. L. (1999). Does parental conflict explain why divorce is negatively associated with child welfare? *Social Forces*, 77, 1283–1316.
- Hawkins, D. N. and Booth, A. (2005). Unhappily ever after: effects of long-term, low-quality marriages on well-being. *Social Forces*, 84, 451–471.
- Härkönen, J. (2014). Divorce: trends, patterns, causes, consequences. In J. K. Treas, J. Scott, and M. Richards (Eds.), *The Wiley-Blackwell Companion to the Sociology of Families* (pp. 303–322). Chichester: John Wiley & Sons.
- Imbens, G. (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 48, 399–423.
- Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62, 467–476.
- Jekielek, S. M. (1998). Parental conflict, marital disruption, and children's emotional well-being. *Social Forces*, 76, 905–935.
- Jonsson, J. O. and Gähler, M. (1997). Family dissolution, family reconstitution, and children's educational careers: Recent evidence for Sweden. *Demography*, 34, 277–293.
- Kim, H. S. (2011). Consequences of parental divorce for child development. *American Sociological Review*, 76, 487–511.
- Lee, D. and McLanahan, S. (2015). Family structure transitions and child development: Instability, selection, and population heterogeneity. *American Sociological Review*, 80, 738–763.
- Mandemakers, J. and Kalmijn, M. (2014). Do mother's and father's education condition the impact of parental divorce on child well-being? *Social Science Research*, 44, 187–199.

- McKinnish, T. G. (2004). Occupation, sex-integration, and divorce. *American Economic Review: Papers and Proceedings*, 94, 322–325.
- McKinnish, T. G. (2007). Sexually integrated workplaces and divorce: Another form of on-the-job search. *Journal of Human Resources*, 42, 331–352.
- McLanahan, S. and Sandefur, G. (1994). *Growing Up with a Single Parent: What Hurts, What Helps?* Cambridge, MA: Harvard University Press.
- McLanahan, S., Tach, L., and Schneider, D. (2013). The causal effects of father absence. *Annual Review of Sociology*, 39, 399–427.
- Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research. Second Edition.* Cambridge: Cambridge University Press.
- Morrison, D. and Cherlin, A. (1995). The divorce process and young children's well-being: A prospective analysis. *Journal of Marriage and Family*, 57, 800–812.
- Ribar, D. C. (2004). What do social scientists know about the benefits of marriage? A review of quantitative methodologies. *IZA Discussion Paper*, No. 998.
- Rijken, A. J. and Liefbroer, A. C. (2012). European views of divorce among parents of young children: Understanding cross-national variation. *Demographic Research*, 27, 25–52.
- Rudolphi, F. (2013). Ever declining inequalities? Transitions to upper secondary and tertiary education in Sweden, for the 1972-1990 birth cohorts. In M. Jackson (Ed.), *Determined to Succeed* (pp. 185–227). Palo Alto, CA: Stanford University Press.
- Sandefur, G. and Wells, T. (1999). Does family structure really influence educational attainment? *Social Science Research*, 28, 331–357.
- Sanz-de-Galdeano, A. and Vuri, D. (2007). Parental divorce and students' performance: Evidence from longitudinal data. *Oxford Bulletin of Economic and Statistics*, 69, 321–338.
- Sigle-Rushton, W., Hobcraft, J., and Kiernan, K. (2005). Parental divorce and subsequent disadvantage: A cross-cohort comparison. *Demography*, 42, 427–446.
- Sigle-Rushton, W., Lyngstad, T. H., Andersen, P. L., and Kravdal, Ø. (2014). Proceed with caution? Parents' union dissolution and children's educational achievement. *Journal of Marriage and Family*, 76, 161–174.
- Small, D. S. and Rosenbaum, P. R. (2008). War and wages: The strength of instrumental variables and their sensitivity to unobserved variables. *Journal of the American Statistical Association*, 103, 924–933.

- Sun, Y. (2001). Family environment and adolescents' well-being before and after parents marital disruption: A longitudinal analysis. *Journal of Marriage and Family*, 63, 697–713.
- Sun, Y. and Li, Y. (2002). Children's well-being during parents marital disruption process: A pooled time-series analysis. *Journal of Marriage and Family*, 64, 472–488.
- Svarer, M. (2007). Working late: Do workplace sex ratios affect partnership formation and dissolution? *Journal of Human Resources*, 42, 583–595.
- Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65, 557–586.
- Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* (pp. 80–108). New York, NY: Cambridge University Press.
- Thomson, E. and Eriksson, H. (2013). Register-based estimates of parents' co-residence in Sweden, 1969-2007. *Demographic Research*, 29, 1153–1186.

**Table A4.1** 2SLS estimates of the effects of parental separation on child GPA within the group of children with low educated parents

	Grade Point Average (GPA)		
	(1) Full sample	(2) Male	(3) Female
Parental separation	-0.32* (0.13)	-0.44* (0.17)	-0.21 (0.18)
Female	0.34** (0.00)		
Birth order	-0.11** (0.00)	-0.10** (0.00)	-0.12** (0.00)
Controls for cohort of child	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	Yes	Yes	Yes
<i>N</i>	142,776	72,295	70,481

*Note:* Second stages of 2SLS IV regression models. Parental separation instrumented by ratio of opposite sex co-workers at the maternal workplace. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

**Table A4.2** Linear Probability Models of the effects of the ratio of opposite sex colleagues at the maternal workplace and the interaction with workplace size on parental separation

	Parental separation					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Ratio opposite sex colleagues at maternal workplace	0.071** (0.003)	0.050** (0.004)	0.050** (0.005)	0.051** (0.005)	0.076** (0.005)	0.029** (0.005)
Workplace size (standardized)	-0.018** (0.001)	-0.007** (0.001)	-0.007** (0.002)	-0.007** (0.002)	-0.008** (0.002)	-0.006** (0.002)
Ratio opposite sex colleagues at maternal workplace X Workplace size (standardized)	0.038** (0.004)	0.019** (0.004)	0.022** (0.006)	0.016** (0.006)	0.026** (0.006)	0.010 (0.005)
Female	0.008** (0.002)	0.007** (0.002)			0.010** (0.003)	0.005* (0.002)
Birth order	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	-0.000 (0.001)	-0.005** (0.001)	0.004** (0.001)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	348,034	348,034	177,831	170,203	142,776	205,258
<i>Cragg-Donald</i>	278.301	108.924	56.366	53.045	101.540	20.096
<i>Wald F statistic</i>						

*Note:* Linear Probability Models. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.

**Table A4.3** Linear Probability Models of the effects of the ratio of opposite sex colleagues in maternal industrial branch on parental separation

	<b>Parental separation</b>					
	(1) Full sample	(2) Full sample	(3) Male	(4) Female	(5) Low parental education	(6) High parental education
Ratio opposite sex colleagues maternal industrial branch	0.077** (0.003)	0.044** (0.004)	0.043** (0.005)	0.045** (0.005)	0.068** (0.006)	0.024** (0.005)
Female	0.008** (0.002)	0.007** (0.002)			0.010** (0.003)	0.004* (0.002)
Birth order	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	-0.000 (0.001)	-0.005** (0.001)	0.004** (0.001)
Controls for cohort of child	Yes	Yes	Yes	Yes	Yes	Yes
Controls for maternal level of education, field of study, workplace size, and region	No	Yes	Yes	Yes	Yes	Yes
<i>N</i>	347,646	347,646	177,623	170,023	142,608	205,038
<i>Cragg-Donald</i>	538.467	142.288	69.907	72.679	146.116	22.549
<i>Wald F statistic</i>						

*Note:* Linear Probability Models. Robust standard errors in parentheses.

Significance levels: \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

*Source:* Sweden in Time - Activities and Relations (STAR) register database.



## **CHAPTER 5: COMPETITION IN THE FAMILY: ESTIMATING AND EXPLAINING EDUCATIONAL INEQUALITIES BETWEEN SIBLINGS<sup>§</sup>**

### **Introduction**

One of the central topics of research on social stratification is the transmission of educational advantage across generations (Breen and Jonsson 2005). Research in this field has mainly concentrated on differences between children belonging to different families. Less attention is paid to inequalities in education which arise between siblings who belong to the same family. However, it is known from research using sibling correlations to measure the impact of family background on educational outcomes that only about half of the variation in education arises between families (Benin and Johnson 1984; De Graaf and Huinink 1992; Hauser and Mossel 1985; Hauser and Wong 1989; Kuo and Hauser 1995; Sieben et al. 2001; Teachman 1995; Toka and Dronkers 1996; Warren et al. 2002). This result implies that the other half of educational inequality is produced within families. In order to provide a complete account how educational inequalities are produced, research on social stratification has therefore to explain why siblings diverge in their educational outcomes.

The analysis of within-family inequality in educational outcomes is also of relevance for research on the intergenerational transmission of education since there are good reasons to expect within-family processes to differ between children with advantaged and children with disadvantaged family backgrounds. Such a hypothesis is supported by theoretical considerations as well as by results of qualitative and quantitative empirical research. Using the notion of “diverging destinies” McLanahan (2004) argues that the demographic processes of the second demographic transition turn out differently in families with many and in families with few resources. In her ethnographic work, Lareau (2011) describes different childrearing practices in lower and upper class families with implications for the educational success of children. DiPrete and Eirich (2006) summarize research on educational inequalities using the concept of “cumulative advantage” which is in line with the idea of an increasing polarization between social classes. Bernardi (2014) argues that the compensatory advantage hypothesis provides a special case of this model. His hypothesis postulates that the negative effects of disadvantageous life events are less negative in families of a higher social origin. More explicitly, using a specific focus on differences between siblings, Conley (2004, 2008)

---

<sup>§</sup> This chapter is under review. The data used in this chapter were made available to me by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin.

puts forward the hypothesis that siblings raised in families of an advantaged social origin are more similar in their educational and occupational outcomes than siblings coming from disadvantaged families.

Despite these advancements, which underline the potential significance of differences between siblings and their variation by social origin for research on the intergenerational transmission of education, research which examines empirically how differences between siblings are brought about and how these processes differ between low- and high-SES families, is still rarely found. There are only two quantitative studies which look at the variation of sibling correlations by family background. Both studies are limited to the United States (Conley and Glauber 2008; Conley et al. 2007). To my mind, this research gap is a major shortcoming because, as sibling correlations demonstrate, a large part of educational inequalities are brought about within and not between families. Insofar as these factors differ between social origin groups, educational inequalities are misjudged by not taking them into account.

This chapter makes two main contributions to an emerging literature. First, I provide descriptive estimates on the variation in the importance of between- and within-family inequalities in educational outcomes for different social groups in Germany. This is the first study to analyze this research question using data on another country than the United States. Second, I test whether the influence of family characteristics, in particular birth order, birth spacing, and maternal age differs with social origin. Previous research has analyzed the influence of these factors on educational outcomes. However, to the best of knowledge, this is the first study which puts the focus on how the impact of these sibling characteristics on child education differs by family socio-economic background.

This chapter proceeds in the following way. First, I summarize the results of previous research on sibling correlations in educational outcomes. Next, I discuss why the similarity of siblings may differ by family socio-economic background. In addition, I discuss why the impact of birth order, birth spacing, and maternal age may differ by social origin. The following section presents the data, variables, and methods used in this paper. Following to that, the results of the analysis are presented. Finally, the last section of this chapter discusses the implications of these results for research on the intergenerational transmission of education.

## **Family Socio-Economic Background and Differences between Siblings in Educational Outcomes**

### Differences and Similarity in Educational Outcomes of Siblings

Both in sociology and in economics, there is an extensive literature which uses sibling correlations in education, occupation, and earnings as a measure of how much family background influences these outcomes. This literature argues that sibling correlations measure the influence of the family on child outcomes (Björklund and Jäntti 2012; Jencks et al. 1972). In addition, sibling correlations capture the influence of other factors which are shared by siblings, e.g. the neighborhood and, often, schools (Jencks et al. 1972; Nicoletti and Rabe 2013; Rabash et al. 2010; Solon et al. 2000). Siblings may also influence each other in their educational outcomes. This sibling influence is also captured in the estimate of sibling similarity (Benin and Johnson 1984; De Graaf and Huinink 1992; Hauser and Wong 1989). For all these reasons, sibling correlations are perceived as measures of the total impact of family background on children. Some studies argue that sibling correlations are lower bound estimates of the impact of family background because there may be factors transmitted from the parents to the children, which are not completely shared among siblings and are not taken into account by sibling correlations (Björklund and Jäntti 2012).

In their influential study, Jencks et al. (1972) analyze sibling correlations in education. This study and subsequent studies on the United States demonstrate that about half of the variance in education can be explained between families (Benin and Johnson 1984; Hauser and Mossel 1985; Hauser and Wong 1989; Kuo and Hauser 1995; Teachman 1995; Warren et al. 2002). Some studies apply this approach to Europe with similar estimates on the importance of inequalities between and within families (De Graaf and Huinink 1992; Sieben et al. 2001; Toka and Dronkers 1996). Comparable to these studies on sibling correlations in education in sociology, several studies in economics, following Solon et al. (1991), report sibling correlations in earnings for the United States (Mazumder 2008), for several Scandinavian countries (Björklund et al. 2002; Björklund and Jäntti 2012), and for Germany (Schnitzlein 2014). Overall, these studies find sibling correlations in earnings to be of a similar size as sibling correlations in education.

Although this research is usually understood as pointing to the importance of family background for educational outcomes of children, a second implication of these results is that siblings differ substantially in their educational (and occupational) outcomes. If half of the

variation in education is produced between families, this implies that the other half is produced within families. For that reason, sociological research should not only focus on analyzing inequalities in education between families but also analyze inequalities in education between siblings.

#### Variation of Sibling Similarity in Education by Family Socio-Economic Background

None of the studies mentioned above analyzed how sibling correlations vary with social origin. This has only been done by two recent papers, which use data on the United States (Conley and Glauber 2008; Conley et al. 2007). There are, however, good reasons to expect the similarity between siblings to vary by family socio-economic background based on both theoretical considerations and empirical research results.

On a theoretical level, economists widely apply Becker and Tomes' (1976) economic theory of the family in order to explain how differences within families occur and discuss whether parents invest similarly in the human capital of all children, whether they invest in order to maximize their returns (Becker 1991; Becker and Tomes 1976), or whether they try to compensate ability differences among their children (Behrman et al. 1982; Griliches 1979). These approaches do mostly not take into account variations by family background although Griliches (1979) points out that compensatory strategies are more feasible for families with many resources.

Therefore, it could be expected that the effects of disadvantageous characteristics or life events vary by family background (Bernardi 2014; Conley 2004, 2008; Hsin 2012). This is because higher social origin families have more resources available to overcome disadvantage. Applied to the setting within families, this implies that lower social origin families reinforce differences between siblings whilst these differences in higher social origin families are balanced out. As a result, we would expect that siblings from higher social origin families resemble each other in their educational outcomes more than siblings from lower social origin families. I label this prediction the compensatory effect of social origin hypothesis.

A similar hypothesis has been advanced by Conley (2004). As stated above, research that tests this hypothesis is scarce. Qualitative evidence supporting this hypothesis is offered by Conley (2004). Conley and Glauber (2008), using data from the US Panel Study of Income Dynamics (PSID), find a lower sibling similarity in years of education for families with a low income and for families with a large family size. They interpret these findings as evidence in

favor of the compensatory effect of social origin hypothesis. However, they do not find significant differences between families with a low and a high level of parental education.

Further support for a perspective that parental investments differ with social origin comes from the literature on the effects of prenatal and early health conditions on long-term outcomes (for an overview see Almond and Mazumder 2013). This literature argues that parents respond to health endowments at birth, with consequences for educational and occupational outcomes. For instance, Almond et al. (2009) use regional variation in Sweden in the prenatal exposure to radiation following the Chernobyl accident in order to analyze the effect of health endowments on child cognitive development. They find that higher radiation negatively influenced the cognitive abilities of children, with the negative effect being concentrated in families in which the father has a low level of education. Similarly, Torche and Echevarría (2011) report that the negative effect of a low birth weight on cognitive development is concentrated in families with low educated mothers. Both Hsin (2012) and Restrepo (2012) analyze how parental investment responses to birth weight differences between their siblings vary with parental education. Both studies come to the conclusion that families with a low level of parental education invest more in children with a higher endowment, while families with a high level of parental education invest more in children with a lower endowment. Bernardi (2014) shows that the negative effect of a young school entry age on grade retention in France is strongly concentrated in lower class families. Ermisch and Francesconi (2013) argue that maternal employment during young childhood has negative effects on the educational outcomes of children from low educated but no negative effects for children from highly educated mothers.

These empirical results from a wide array of different studies are in line with the compensatory effect of social origin hypothesis. They also support the notion that characteristics which vary between siblings and which influence child development, lead to differential treatment by social origin and, as a consequence, to differences in educational outcomes between siblings. The next section discusses the three factors that I will address in this paper and the expectations of how and why the impact of these factors may differ by family socio-economic background.

## Factors which Influence Educational Inequalities between Siblings and their Variation by Family Socio-Economic Background

There has been some research on how family characteristics influence educational outcomes. The possible influence of family size on child education has received a lot of attention in research on educational inequalities (e.g. Blake 1989; Downey 1995; Guo and VanWey 1999). This chapter, however, concentrates on those factors that can explain inequalities between siblings: birth order, birth spacing, and maternal age. Family size is not included in this analysis because it does not vary within families and, hence, cannot explain differences between siblings.

In particular, I am interested in whether the influence of birth order, birth spacing, and maternal age varies by family background. The compensatory effect of social origin hypothesis outlined above leads to the expectation that the negative effects of disadvantageous life events may have stronger consequences for educational outcomes in lower than in higher social origin families. This is because, according to the argument advanced here, parents in lower social origin families reinforce ability differences between their children whilst parents in higher social origin families tend to balance them out (Conley 2004).

Researchers have been analyzing birth order effects for a long time. There are now several studies which demonstrate that birth order influences cognitive abilities as well as educational outcomes (Barclay 2014; Bjerkedal et al. 2007; Black et al. 2005; Booth and Kee 2009; Härkönen 2014; Kristensen and Bjerkedal 2007). Black et al. (2005) report that birth order effects dominate family size effects in Norway. Other studies find similar results for the United States (Booth and Kee 2009). There is evidence from family-fixed effects models that birth order has an impact on educational transitions in Germany. Härkönen (2014) reports that second-born children have an around four percentage points lower probability of graduating from the highest educational track (*Gymnasium*) in the German education system than first-borns.

Previous research does not report differences by family socio-economic background in the effects of birth order on education (Black et al. 2005; Härkönen 2014). In so far as birth order effects occur through financial resource limitations, we would expect them to be stronger in lower social origin families. It could, however, be that the underlying mechanism of birth order effects are parental time restrictions, which should apply similarly to both lower

and higher social origin families. In the latter case, we would expect no variation in birth order effects between social origin groups.

The second factor, which I include, is birth spacing. In two papers, Powell and Steelman (1990, 1993) argue that having more closely spaced siblings is detrimental to educational outcomes. Contrary to that, using data on Hungary van Eijck and de Graaf (1995) find positive instead of negative effects of having a closely spaced sibling on educational outcomes. One difference between their study and Powell and Steelman's (1990, 1993) work is that they use an interval of six years to define closely spaced siblings, instead of a two year interval. Using exogenous variation in child spacing due to a Swedish reform, Pettersson-Lidbom and Skogman Thoursie (2009) estimate close birth spacing to have negative consequences for university attendance. Using miscarriages between live births as an instrumental variable for birth spacing and data on the United States, Buckles and Munnich (2012) support the notion that close birth spacing has a negative effect on educational performance.

None of these studies investigates whether the negative effects of having more closely spaced siblings differs by family background. However, if the dilution of parental resources is the main mechanism underlying the negative effects of having more closely spaced siblings on education, as argued by Powell and Steelman (1995), I expect the negative effect of having more closely spaced siblings on educational attainment to be more pronounced in lower social origin families.

As a third factor I include parental age in the form of maternal age in the analysis. Mare and Tzeng (1989) report positive effects of a higher paternal age on educational attainment of sons in the United States. Using data on the Netherlands, Kalmijn and Kraaykamp (2005) demonstrate that these positive effects of a higher maternal (and paternal) age on educational outcomes are also brought about using family-fixed effects models. They find positive effects of a higher maternal age on years of education.<sup>1</sup>

These studies do not test whether the positive effects of a higher maternal age, or, in other words, the negative effects of a younger maternal age vary with social origin. The underlying mechanism of the maternal age effect may be, as argued by Kalmijn and Kraaykamp (2005), an increase in a mother's (and possibly also a father's) skills over the life course. In line with this hypothesis, Powell et al. (2006) show that a higher parental age is associated with the transmission of more economic, cultural, and social resources to the child. The negative effect of a young maternal age may be stronger in lower social origin families for two reasons. First, young low-SES mothers and fathers may have particularly few skills

and resources. They may then accumulate relatively more skills and resources over the life course than high-SES parents. Second, skills and resources may have fewer consequences for children's educational outcomes in higher social origin families since they are more likely to be able to compensate the lack of skills and resources of one type through skills and resources of other types.

## **Data and Methods**

### Data and Sample Selection

The analysis in this paper uses data on a sample of 1,999 siblings from 900 families, derived from the German Socio-Economic Panel Study (SOEP). The SOEP is a long-running panel study which samples household units (Wagner et al. 2007). The study started in 1984 with a sample of West German households and includes a sample of East German households from 1991 onwards. Since all household members are sampled and observed after they leave their initial household, a sample of siblings born into the same households can be constructed.

The sample is restricted to respondents who filled out a special questionnaire in the year in which they turned 17 years, which provides information on educational outcomes in the year preceding the survey, i.e. when most of the respondents were 16 years old. The response rate of this youth questionnaire was 84.4% in 2012 (TNS Infratest Sozialforschung 2013). The sample includes all respondents who were born in SOEP households between 1982 and 1995. Respondents without any siblings with valid information are dropped from the analyzed sample.<sup>2</sup> In addition, I dropped children from the sample who experienced the death of one of their parents during childhood.

The sample includes children born in East Germany, West Germany, and abroad. In one part of the analysis, I analyze sibling correlations separately by migration background. Someone is defined as having a migration background if both parents were born outside of Germany. I conduct the analysis for children from both parts of Germany together. These children have experienced most of their childhood in reunified Germany and all school results were obtained after reunification. However, I report sibling correlations separately for East and West Germany in one part of the analysis. The East German sample is a sample of respondents whose parents lived in the German Democratic Republic (GDR) in 1989. Similarly, in the West German sample all children are included whose parents lived in the Federal Republic of Germany (FRG) in 1989.

The analytical sample in order to analyze cognitive performance as an outcome variable is reduced in size since the test of cognitive performance was only introduced in 2006 and, hence, information on cognitive performance is only available for children born between 1987 and 1995. In order to test the robustness of the results for the other two outcome variables, I estimated these models on the reduced sample which only includes respondents with valid information on their cognitive performance. These models led to nearly identical results as I obtained for the larger sample (results available upon request).

### Variables and Descriptive Statistics

I use three measures of educational outcomes: a measure of cognitive performance, a measure of Grade Point Average (GPA) at age 16 years based on school grades in German and Mathematics, which are the two main subjects in the German education system, and a dichotomous variable, which indicates whether someone attends the highest track in secondary school (*Gymnasium*) leading to *Abitur* (the highest school degree in Germany that is required to access university, comparable to *A Level* in the UK).

The measure of cognitive performance is based on a cognitive test included in the survey since 2006. The aim of this test is to measure fluid cognitive performance. The three parts of the test measure the ability to perform word analogies, numeric cognitive potential, and figural cognitive potential. A detailed account of the questions used to measure these three components is provided in Schupp and Herrmann (2009). As an outcome variable I use a measure of all correct answers in all three parts of the test. I standardize this variable to have a mean of 0 and a standard deviation of 1. Since this test was only conducted from 2006 onwards, the analysis sample of this outcome is restricted to respondents born between 1987 and 1995. For that reason, the sample used to estimate this outcome includes only about half of the respondents in the samples used to estimate the other two outcomes.

I use information on school grades at age 16 years to construct a measure of Grade Point Average (GPA). In Germany, school grades range from 1 to 6, 1 indicating the best possible grade. I recode the variables on grades in German and Mathematics so that a higher number signifies a better grade. In this case, a positive effect of an independent variable can be understood as an effect that increases educational performance. I take the average between both grades in order to obtain a measure of Grade Point Average (GPA). I standardize the variable so that it has a mean of 0 and a standard deviation of 1. Since these grades are achieved in different school tracks, I control for track attendance in the family-fixed effects

models that predict GPA. The fact that grades are obtained in different tracks explains also the lower sibling correlations for GPA than for cognitive abilities and track attendance.

Research has argued that early tracking, which happens when the children are 10 to 12 years old, is the main reason for high inequality in educational outcomes in Germany (Hillmert and Jacob 2010). Consequently, attendance of the upper track can be used as a proxy for final educational attainment. Track attendance is a dichotomous variable, which is coded 1 for all respondents who attend the highest track in the German education system at age 16 to 17 years and 0 for all other respondents. This latter group includes high school dropouts and those who attend lower tracks. Respondents who attend so-called comprehensive schools (*Gemeinschaftsschulen*), which combine all tracks into one school, are dropped from the analysis of this outcome. The track that these respondents attend within a comprehensive school cannot be determined. Since this is the case for only 128 children, it seems unlikely that this reduction in the analyzed sample induces any bias.

In the second part of the analysis, I use family-fixed effects models in order to test the impact of birth order, birth spacing, and maternal age on education. Birth order is a continuous variable, which gives the rank within the sibling group. Birth spacing is coded by counting the number of siblings born within an interval of two years around a respondent's birth. In other words, this variable counts the number of siblings born the previous, the same, or the year after the respondent. Maternal age is a continuous variable reporting maternal age at birth.

I measure social origin via three indicators, the first one being parental education. Parental education is measured by the highest level of education achieved by either parent. I employ a dummy variable, which is coded 1 if one or both parents have received *Abitur* (or an equivalent qualification) and 0 otherwise. In addition, I have employed parental EGP class and parental ISEI as alternative measures of social origin. High parental class is defined as a dummy variable, which is set to 1 if one of the parents belongs to one of the service classes or is self-employed. Parental ISEI is a continuous variable. In order to compare two groups I assign those who have a parental ISEI score higher than 50 as being of a high parental ISEI origin and those with a score of 50 or lower as being of a low parental ISEI origin.

I control in all models for gender with a dummy variable, which is coded 1 for male respondents. In addition, in the models predicting GPA I control for track attendance because grades are given to children attending different tracks in the German education system.

Descriptive statistics on the sample used in the analysis are reported in Table 5.1. Since there are only few missing values on the outcome variables and virtually no missing values on the independent variables, I do not impute missing values.<sup>3</sup>

Table 5.1 also reports the decomposition of the variance of the variables into a between- and a within-family component. This decomposition demonstrates that, although most of the variation of the variables is between families, there is also considerable variation within families. With respect to all explanatory and outcome variables, the within-family standard deviation is sufficiently different from 0 which should allow me to discover effects if the underlying effects are large enough.

**Table 5.1** Descriptive statistics

<b>Variable</b>	<b>Mean</b>	<b>Std. dev.</b>	<b>Std. dev. between families</b>	<b>Std. dev. within families</b>	<b>Min.</b>	<b>Max.</b>	<b>N</b>
Number of children in data <sup>1</sup>	2.34	0.64	-	-	2	6	1,999
Age in 2012	23.43	3.36	2.78	1.99	17	29	1,999
Cognitive performance <sup>2</sup>	0	1	0.91	0.48	-2.87	2.54	1,060
Grade Point Average (GPA)	0	1	0.77	0.65	-3.98	2.52	1,976
Upper track attendance ( <i>Gymnasium</i> ) <sup>3</sup>	0.40	0.49	0.43	0.24	0	1	1,818
High parental education	0.34	0.48	-	-	0	1	1,999
High parental class	0.48	0.50	-	-	0	1	1,961
Parental ISEI	49.46	17.75	-	-	16	90	1,883
Migration background	0.23	0.42	-	-	0	1	1,984
East Germany (GDR origin)	0.23	0.42	-	-	0	1	1,904
Male	0.50	0.50	0.34	0.37	0	1	1,999
Birth order	1.91	0.99	0.69	0.65	1	10	1,999
Number of closely spaced siblings	0.24	0.46	0.42	0.19	0	3	1,999
Maternal age at birth	33.63	4.70	4.72	0.43	18	52	1,999

*Note:*

<sup>1</sup> This variable counts the number of children from the respondent's family in the data and includes, hence, the respondent and his siblings.

<sup>2</sup> The cognitive performance test was only introduced in 2006. For that reason only children born between 1987 and 1995 participated in it.

<sup>3</sup> Those who attend comprehensive schools (*Gemeinschaftsschulen*) are dropped from the analysis of this outcome because their track cannot be determined.

*Source:* German Socio-Economic Panel Study (SOEP), v29.

## Analytical Strategy

In order to achieve the two aims of this paper I employ two techniques. First, I use multilevel models to estimate sibling correlations in educational outcomes. Second, I employ family-fixed effects models to test the impact of birth order, birth spacing, and maternal age on educational outcomes.

First, I provide descriptive estimates of the role inequalities within families play in bringing about educational inequalities. For this purpose I estimate restricted maximum likelihood multilevel models and report the intra-class correlation coefficients (ICCs) of these models (Björklund and Jäntti 2012; Mazumder 2008; Schnitzlein 2014). This approach is comparable, although not identical in the technical aspects, to the one taken by Conley and Glauber (2008), following the work of Solon et al. (1991). The analysis of multilevel models restricted to specific social groups tests whether sibling similarity varies by family socio-economic background (Conley et al. 2007; Conley and Glauber 2008).

In these models I only control for year of birth and gender, as this has been done in previous research. Hence, these multilevel models can be written as (cf. Schnitzlein 2014):

$$E_{i,j} = \mathbf{X}_{i,j} \beta + \alpha_j + \delta_{i,j} \quad (1)$$

with  $E_{i,j}$  being the educational outcome of interest,  $\mathbf{X}_{i,j}$  being a vector of the birth year dummy and gender control variables,  $\alpha_j$  being the family-specific, and  $\delta_{i,j}$  being the individual-specific component.

From these models I report and compare the intra-class correlation coefficients  $\rho$ :

$$\rho = \sigma_{\alpha}^2 / (\sigma_{\alpha}^2 + \sigma_{\delta}^2) \quad (2)$$

with  $\sigma_{\alpha}^2$  being the variance of the family-specific and  $\sigma_{\delta}^2$  being the variance of the individual-specific component. Hence, the ratio  $\rho$  reports how much of the total variance  $\sigma_{\alpha}^2 + \sigma_{\delta}^2$  is shared by siblings. To give an example for Germany, Schnitzlein (2014) reports an intra-class correlation  $\rho$  in earnings of 0.432 for brother and 0.391 for sisters.

Second, I test the influence of certain individual-level mechanisms on differences between siblings. For this purpose I employ so-called family-fixed effects (or sibling-difference) models. Family-fixed effects models are similar to individual-fixed effects models with the difference being that the outcomes of an individual are compared not over several time points but to siblings who have grown up in the same family.

The family-fixed effects model can be written in the following way (cf. Conley et al. 2007):

$$Y_{ij} - Y_{ik} = \beta (\mathbf{X}_{ij} - \mathbf{X}_{ik}) + (\alpha_i - \alpha_i) + (\mu_{ij} - \mu_{ik}) \quad (3)$$

with  $Y$  being the educational outcome of interest, the subscript  $i$  refers to the family with  $j$  and  $k$  referring to two siblings from that family.  $X$  is a vector of explanatory variables (in this chapter birth order, birth spacing, and maternal age). The error component of this model consists of a family-specific error term and an individual-specific error term. The model controls for the family-specific error term. The estimated models in the analysis allow for that there can be more than two siblings in each family and that the number of siblings per family varies. In families with more than two children the information on all siblings is taken into account.

Not only does employing family-fixed effects models allow me to control for unobserved heterogeneity, it is also the only approach that makes it possible to analyze how inequalities *within* families occur. Only differences between siblings of the same family contribute to the coefficients of these regression models. I interact the measures of birth order, birth spacing, and maternal age with parental education in order to see whether the influence of birth order, birth spacing, and maternal age varies by family socio-economic background.

There are limitations in the application of family-fixed effects models. Frisell et al. (2012) discuss that the omission of variables that vary between siblings and measurement error can lead to bias in family-fixed effects models. It is, hence, important to control for factors that vary between siblings, such as gender. There may still be bias through confounding variables which lead to differences in educational outcomes between siblings but are not included as control variables. There is, however, little research on which control variables need to be applied in family-fixed effects models and how they correlate with birth order, birth spacing, and maternal age. For that reason, I do not expect any remaining bias to be very large in size. The problem of measurement error is reduced in this paper by using data from a high quality panel study. In addition, reverse causality could be a problem in the sense that these models do not take into account parental fertility responses to children's endowments (Ermisch and Francesconi 2013). Similar to the case of control variables in family-fixed effects models, this is a topic on which more empirical research is needed. I am not able to do this in this paper and assume that any effect of reverse causality is small. In any case, we should be aware of that causal interpretations of the family-fixed effects models rest on two assumptions: First, that all confounding variables which vary between siblings are controlled for and, second, that there is no reverse causality in the sense that children's endowments influence parental fertility decisions.

Cognitive performance and school grades are standardized continuous variables which are estimated using OLS regression. Upper track attendance is a dichotomous outcome variable

which is estimated using Linear Probability Models. The use of Linear Probability Models allows me an easy interpretation of regression coefficients, in particular, it makes it possible to compare coefficients across models without them being biased through unobserved heterogeneity (Angrist and Pischke 2009; Mood 2010). In the models that I estimate there are virtually no out-of-sample predictions (between 0 and 0.3 percent), so this should not be a reason to turn to logit and probit models which are based on stronger assumptions and more difficult to interpret.

The multilevel models are estimated using the `xtmixed` command in STATA 13.1. Family-fixed effects models are estimated using the `xtreg` command in the same program. Since the direction of the influencing factors is clearly predicted by theoretical expectations and empirical results of previous research, all significance tests in the family-fixed effects models are based on one-tailed t-tests (Freedman et al. 2007).

## **Results**

### The Stratification of Sibling Similarity in Educational Outcomes

Table 5.2 reports intra-class correlation coefficients from multilevel regression models which are called sibling correlations. A higher correlation implies that siblings are more similar and, therefore, that more inequality is produced between as opposed to within families. The models are estimated separately for the three analyzed outcomes. Results are reported for the overall sample and then for samples restricted to social groups. The main aim of this exercise is to give a picture of how the relation between inequalities between and within families varies by family socio-economic background.

The overall results, reported in the first row of Table 5.2, suggest an important role for inequalities within families. These results underline the primary proposition of this paper; there is considerable variation between siblings, which has to be explained. Most variance is explained within rather than between families. This is demonstrated by low intra-class correlations. This conclusion particularly holds for measures of performance. The intra-class correlation for track attendance is higher. Hence, differences are stronger concentrated between than within families for track attendance than for cognitive performance. This means that educational attainment is more unequally distributed between than within families than educational performance.<sup>4</sup>

With respect to the stratification of sibling similarity in education, the results do not reveal a clear pattern. Three indicators are used, which can be interpreted as proxies for social origin: parental education, parental class, and parental ISEI. Independent of which of these indicators is used to define social origin, the coefficients and, hence, the relations between the variation within and between families are mostly similar across both social origin groups. With respect to GPA, however, there is a clear pattern of stratification with a higher sibling similarity in high- than in low-SES families. Nevertheless, this finding is not reproduced using the other two measures of child education. Since there are no systematically higher sibling correlations for families with a higher social origin, there is no general support for the hypothesis that children from lower social origin families show a higher variation in educational outcomes.

**Table 5.2** Sibling correlations in educational outcomes

	<b>Cognitive performance</b>	<b>Grade Point Average (GPA)</b>	<b>Upper track attendance (Gymnasium)</b>
Overall sample ( <i>N</i> = 1,999)	0.46 (0.04)	0.22 (0.03)	0.53 (0.02)
High parental education ( <i>N</i> = 687)	0.36 (0.07)	0.29 (0.05)	0.40 (0.05)
Low parental education ( <i>N</i> = 1,312)	0.41 (0.05)	0.16 (0.04)	0.43 (0.03)
High parental class ( <i>N</i> = 949)	0.30 (0.06)	0.24 (0.04)	0.44 (0.04)
Low parental class ( <i>N</i> = 1,012)	0.47 (0.06)	0.17 (0.04)	0.44 (0.04)
High parental ISEI ( <i>N</i> = 1,023)	0.40 (0.05)	0.26 (0.04)	0.47 (0.04)
Low parental ISEI ( <i>N</i> = 976)	0.41 (0.06)	0.15 (0.04)	0.43 (0.04)
No migrant background ( <i>N</i> = 1,525)	0.41 (0.05)	0.20 (0.03)	0.54 (0.03)
Migrant background ( <i>N</i> = 459)	0.38 (0.09)	0.25 (0.06)	0.42 (0.06)
West Germany (FRG origin) ( <i>N</i> = 1,463)	0.43 (0.04)	0.22 (0.03)	0.56 (0.03)
East Germany (GDR origin) ( <i>N</i> = 441)	0.65 (0.07)	0.20 (0.06)	0.45 (0.06)

*Note:* Tables report intra-class correlation coefficients of multilevel models estimated with restricted maximum likelihood. Standard errors in parentheses.

*Source:* German Socio-Economic Panel Study (SOEP), v29.

One further difference may arise between families with and without a migration background. Comparing these two groups, sibling correlations are for two out of three

outcomes higher in non-migrant families. This finding can be interpreted as supporting the compensatory effect of social origin hypothesis with respect to migrants.

Finally, the last two rows compare the outcomes between families who lived in the Federal Republic of Germany in 1989, called West German families, and families that lived in the German Democratic Republic, called East German families. Sibling correlations are similar for these groups. This suggests that processes of intergenerational transmission of education are comparable for East and West Germans.

To sum up, this analysis has generally provided no support for a stratification of sibling similarity in educational outcomes. Contrary to the theoretical expectations, sibling correlations are not systematically higher among high- than among low-SES families. The only exceptions to this conclusion are the findings with respect to school grades. Sibling correlations in GPA are indeed higher in high- than in low-SES families.

#### Explaining Differences between Siblings in Educational Outcomes

Sibling correlations are informative. But if inequality between siblings is largely caused by locating one social origin group at the upper end of an outcome variable and another group at the lower end, sibling correlations within these groups can be of a similar size while at the same time the meaning and the consequences of the similarity between siblings can be very different between social origin groups. For instance, the above results demonstrate that sibling correlations in upper track attendance are of a similar size for families with higher and lower educated parents. However, this similarity takes place at different levels. Most of the children with parents with a high level of education attend the upper track in high school (*Gymnasium*), while most of the children from families with a low level of parental education attend one of the lower tracks. The meaning of a similar sibling correlation in track attendance is very different between lower and higher social origin families.

The question that arises is whether there is stratification in the impact of influencing factors on educational outcomes. It may be that children from higher social origin families attend the upper track in any case, but children from lower social origin families need favorable circumstances to do so. In that case, the factors that influence educational outcomes at the individual level and, hence, the factors that cause differences between siblings may differ between children with different socio-economic backgrounds.

In order to test this kind of stratification in the impact of family characteristics on education, I run family-fixed effects models which include the interactions between the

influencing characteristics and social origin. In Table 5.3, I present results using parental education as an indicator of social origin. As a robustness check, I use parental class to operationalize social origin.

**Table 5.3** Family-fixed effects models of the impact of differences between siblings on educational outcomes

	Cognitive performance <sup>1</sup>		Grade Point Average (GPA) <sup>1</sup>		Upper track attendance (Gymnasium) <sup>2</sup>	
	(1)	(2)	(3)	(4)	(5)	(6)
Birth order	-0.12** (0.04)	-0.15** (0.06)	-0.07* (0.03)	-0.10* (0.04)	-0.03** (0.01)	-0.02† (0.01)
Number of closely spaced siblings	0.10 (0.15)	0.14 (0.20)	-0.15† (0.11)	-0.03 (0.14)	-0.03 (0.05)	-0.08† (0.05)
Maternal age	0.12† (0.08)	0.02 (0.09)	0.03 (0.05)	-0.02 (0.02)	-0.00 (0.00)	0.00 (0.01)
Male	0.16* (0.07)	0.15* (0.07)	-0.22** (0.06)	-0.23** (0.06)	-0.06** (0.02)	-0.06** (0.02)
Birth order X High parental education		0.07 (0.08)		0.06 (0.06)		-0.02 (0.03)
Number of closely spaced siblings X High parental education		-0.08 (0.30)		-0.27 (0.21)		0.14† (0.11)
Maternal age X High parental education		0.18* (0.09)		0.19** (0.08)		0.00 (0.01)
<i>N</i>	1,060	1,060	1,925	1,925	1,818	1,818

*Note:* Cluster-robust standard errors in parentheses.

Significance levels: †  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ .

<sup>1</sup> OLS Regression Models. Outcome variables are standardized with a mean of 0 and a standard deviation of 1.

<sup>2</sup> Linear Probability Models.

*Source:* German Socio-Economic Panel Study (SOEP), v29.

Model 1 and 2, which are reported in Table 5.3, reveal, as expected, a negative effect of birth order on cognitive performance. In addition, I find the expected positive effect of a higher maternal age on cognitive performance. I cannot find any negative effect of having more closely spaced siblings on this outcome variable. The analysis of cognitive performance also reveals some differences in the impact of these sibling characteristics in families with a higher and a lower level of parental education. The positive effect of a higher maternal age on cognitive performance is concentrated in families with a high level of parental education. The effect size is of substantive importance. Having a five years older mother, which is about one

standard deviation of the distribution of maternal age in the sample, leads to an increase in cognitive performance by around one standard deviation in families with a high level of parental education. By contrast, there is no effect of maternal age on cognitive performance in families with a low level of parental education.

Model 3 and 4 in Table 5.3 report the effects of birth order, birth spacing, and maternal age on GPA. Overall, these results are similar to the results for cognitive performance. Birth order has the expected negative impact on school grades, which does not vary by social origin. In line with the results on cognitive performance, the positive effect of a higher maternal age is concentrated in families with a high level of parental education. In addition, there is a negative effect of the number of closely spaced siblings in both families with a high and with a low level of parental education. This was not the case for cognitive performance. The effect, however, is small. Having one additional closely spaced sibling leads to an approximately 1/5 standard deviations lower GPA. There is no sizeable variation in the birth spacing effect by family socio-economic background.

In contrast to cognitive performance, GPA is an educational outcome that parents can directly observe. School grades are, however, not as important as a predictor of final educational attainment as track attendance. The final two models in Table 5.3 analyze the effects of birth order, birth spacing, and maternal age on track attendance, which is arguably the most important outcome in the German education system. In particular, track attendance is most influential for future educational success (Hillmert and Jacob 2010).

Model 6 demonstrates that the negative effect of having more closely spaced siblings is concentrated in families with a low level of parental education. Having more closely spaced siblings has a negative effect on upper track attendance only for children from lower social origin families, even though no influence of close birth spacing on cognitive performance and GPA was observed for these children. This finding is in line with the resource dilution explanation of the effect of birth spacing on education (Powell and Steelman 1993, 1995). In families with a low level of parental education, the probability of attending the upper track is reduced by around eight percentage points for each additional closely spaced sibling. This effect is of a substantive large size and larger than the male penalty, which is only around six percentage points.

Neither the effect of birth order nor the effect of maternal age on track attendance varies by social origin. Birth order has a negative effect on track attendance, which is of similar size for both social origin groups. Maternal age does not affect track attendance. This implies that

the positive effects of maternal age on educational performance in families with a high level of parental education do not have consequences for track attendance.

### Robustness Checks

I conducted several robustness checks in order to test whether the results of the family-fixed effects models are robust to using another indicator of social origin. Table 5.4 reports results if parental class is used to operationalize social origin instead of parental education.

With respect to the effect of the number of closely spaced siblings, the previous results are confirmed in this robustness check. The interaction effect between birth spacing and parental class is statistically significant and the effect is of a similar size than in the specification which uses parental education as a measure of social origin. For each additional closely spaced sibling, the probability to attend the upper track is reduced by about ten percentage points in lower class families. In upper class families there is no negative effect of having more closely spaced siblings on track attendance.

In addition, this robustness check confirms the result that the effects of birth order do not vary by family background. There are no class differences in birth order effects for any of the three educational outcomes. The same result has been produced in the analysis which used parental education as a measure of social origin. Hence, this finding holds up when different indicators of social origin are used.

Contrary to the findings for birth spacing and birth order, the result for the effect of maternal age is not robust to operationalizing social origin via parental class. According to the model predicting cognitive performance, the positive effect of a higher maternal age is concentrated among lower social class families. Since I do not have a good explanation why this effect should be stronger in higher educated (see previous section) *and* lower class families, I conclude that the result is not robust to different specifications of social origin and should, therefore, be interpreted with caution.

**Table 5.4** Family-fixed effects models of the impact of differences between siblings on educational outcomes, robustness check with parental class

	Cognitive performance <sup>1</sup>		Grade Point Average (GPA) <sup>1</sup>		Upper track attendance (Gymnasium) <sup>2</sup>	
	(1)	(2)	(3)	(4)	(5)	(6)
Birth order	-0.12** (0.04)	-0.13* (0.06)	-0.08* (0.03)	-0.07† (0.04)	-0.03** (0.01)	-0.03† (0.02)
Number of closely spaced siblings	0.09 (0.16)	0.09 (0.18)	-0.14† (0.11)	-0.03 (0.14)	-0.03 (0.05)	-0.11* (0.06)
Maternal age	0.12† (0.08)	0.14† (0.09)	0.03 (0.05)	0.04 (0.06)	-0.00 (0.00)	-0.00† (0.00)
Male	0.16* (0.07)	0.15* (0.07)	-0.21** (0.06)	-0.21** (0.05)	-0.06** (0.02)	-0.06** (0.02)
Birth order X High parental class		0.03 (0.09)		-0.01 (0.07)		-0.01 (0.03)
Number of closely spaced siblings X High parental class		0.01 (0.34)		-0.24 (0.21)		0.16† (0.10)
Maternal age X High parental class		-0.27** (0.09)		-0.06 (0.12)		0.04 (0.03)
<i>N</i>	1,046	1,046	1,889	1,889	1,789	1,789

*Note:* Cluster-robust standard errors in parentheses.

Significance levels: †  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ .

<sup>1</sup> OLS Regression Models. Outcome variables are standardized with a mean of 0 and a standard deviation of 1.

<sup>2</sup> Linear Probability Models.

*Source:* German Socio-Economic Panel Study (SOEP), v29.

## Discussion and Conclusion

This study has both provided descriptive figures of the importance of inequalities between and within families across social origin groups and has analyzed which factors cause differences between siblings in educational outcomes. There is a substantial part of inequality that exists between siblings. The analysis reveals that the amount of sibling similarity in educational outcomes can vary by family socio-economic background. However, this variation does not constantly show the systematic pattern predicted by the compensatory effect of social origin hypothesis and found in previous research on the US (Conley et al. 2007; Conley and Glauber 2008).

The analysis has also shown that the impact of some family characteristics on child education differs by family socio-economic background. The negative effects of having more

closely spaced siblings are only found in families with a low level of parental education. Contrary to that, the negative effects of a younger maternal age do not vary between social origin groups in a systematic way. In addition, the effects of birth order do not differ significantly between social origin groups.

These results add close birth spacing to a list of disadvantageous family characteristics and life events which have reduced negative effects on educational outcomes in higher social origin families. Previous research has obtained a similar finding of a compensatory effect of social origin with respect to the negative effects on educational outcomes of shocks in early health conditions (Almond et al. 2009), a low birth weight (Torche and Echevarría 2010), maternal employment (Ermisch and Francesconi 2013), and a young school entry age (Bernardi 2014).

Several underlying mechanisms can possibly cause the observed interaction between social origin and factors that negatively influence educational outcomes. On the one hand, a compensatory behavior among higher social origin families can bring about the compensatory effect. On the other hand, these results can also be caused by a reinforcing behavior among lower social origin families.<sup>5</sup> Testing which of these behavioral mechanisms causes social origin differences in the effects of these characteristics on educational outcomes is beyond the scope of this chapter but should be a major objective of further research.

This chapter, however, also demonstrates that a compensatory effect of social origin is not observed with respect to every life event or characteristic which negatively influences educational outcomes. Birth order effects are quite similar across social origin groups; the negative effects of a younger maternal age on child outcomes are stronger in families with a high level of parental education but weaker in higher class families. These findings deserve special attention of further research on the compensatory effect of social origin because they may shed light on the mechanisms behind the compensatory effect. A general theory of the compensatory effect which describes the mechanisms may, hence, investigate why in some instances no compensatory effect emerges.

The results presented here have significance for research on educational inequalities in general. They show that what happens within families has to be taken into account in estimating and explaining the intergenerational transmission of education. Similarities between siblings are lower than theories of social stratification predict. The integration of differences within families in theories of educational and social mobility is still an open challenge. It is not an easy but a necessary one; in particular, the effects of different processes

between versus within families are both in their own right an important part of understanding educational inequalities.

This conclusion also implies that sibling correlations have to be interpreted with caution. They do not take into account differences in the way inequalities within families occur between social origin groups. If, as I have shown, the effects of characteristics that influence educational attainment within families vary between social origin groups, sibling correlations report biased estimates of the overall effects of family background on educational and occupational outcomes (Hsin 2012).

In general, the findings of this study underscore the need for further research on siblings. The increasing availability of data sources with information on siblings will make this a feasible enterprise. What we can hope for in using these data is obtaining a more complete picture of how educational inequalities occur.

#### ENDNOTES

1. As noted by both Mare and Tzeng (1989) and Kalmijn and Kraaykamp (2005) paternal and maternal age are highly correlated, not allowing them to separate whether the effects of parental age are due to the father's or mother's age. Nor can I distinguish whether the effects are due to maternal or paternal age. Therefore I only report results using maternal age.
2. There is also a small number of children who experienced the death of a parent during childhood. These children were dropped from the analyzed sample because the death of a parent may have an influence on educational outcomes.
3. The only variable for which there is a large number of missing values is parental ISEI. However, this variable is not used in the second part of the analysis.
4. This conclusion is mainly based on the comparison of sibling correlations in cognitive performance and track attendance. Sibling correlations in GPA are lower than sibling correlations in cognitive performance and track attendance since they are obtained in different school tracks.
5. With respect to birth weight, both Hsin (2012) and Restrepo (2012) show that the social origin of the family influences the way parents respond to birth weight differences. However, since they do not include educational outcomes in their analyses, both studies cannot show that these diverging parental investment behaviors explain social origin differences in educational outcomes.

## REFERENCES

- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124, 1729–1772.
- Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annual Review of Economics*, 5, 37–56.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Barclay, K. (2014). Birth order and educational attainment: Evidence from fully adopted sibling groups. *Intelligence*, 48, 109–122.
- Becker, G. S. (1991). *A Treatise on the Family. Enlarged Edition*. Cambridge, MA: Harvard University Press.
- Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84, S143–S162.
- Behrman, J. R., Pollak, R. A., and Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, 90, 52–73.
- Benin, M. H. and Johnson, D. R. (1984). Sibling similarities in educational attainment: A comparison of like-sex and cross-sex sibling pairs. *Sociology of Education*, 57, 11–21.
- Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: a regression discontinuity based on month of birth. *Sociology of Education*, 87, 74–88.
- Bjerkedal, T., Kristensen, P., Skjeret, G. A., and Brevik, J. I. (2007). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence*, 35, 503–514.
- Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., and Österbacka, E. (2002). Brother correlations in earnings in Denmark, Finland, Norway and Sweden compared to the United States. *Journal of Population Economics*, 15, 757–772.
- Björklund, A. and Jäntti, M. (2012). How important is family background for labor-economic outcomes? *Labour Economics*, 19, 465–474.
- Black, S. E., Devreux, P. J., and Salvanes, K. G. (2005). The more, the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, 120, 669–770.
- Blake, J. (1989). *Family Size and Achievement*. Berkeley, CA: University of California Press.

- Booth, A. L. and Kee, H. J. (2009). Birth order matters: The effect of family size and birth order on educational attainment. *Journal of Population Economics*, 22, 367–397.
- Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31, 223–243.
- Buckles, K. S. and Munnich, E. L. (2012). Birth spacing and sibling outcomes. *Journal of Human Resources*, 47, 613–642.
- Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York, NY: Pantheon Books.
- Conley, D. (2008). Bringing sibling differences in: Enlarging our understanding of transmission of advantage in families. In D. Conley and A. Lareau (Eds.), *Social Class: How Does It Work?* (pp. 179–200). New York, NY: Russell Sage Foundation.
- Conley, D. and Glauber, R. (2008). All in the family? Family composition, resources, and sibling similarity in socioeconomic status. *Research in Social Stratification and Mobility*, 26, 297–306.
- Conley, D., Pfeiffer, K. M., and Velez, M. (2007). Explaining sibling differences in achievement and behavioral outcomes: The importance of within- and between-family factors. *Social Science Research*, 36, 1087–1104.
- De Graaf, P. and Huinink, J. (1992). Trends in measured and unmeasured effects of family background on educational attainment and occupational status in the Federal Republic of Germany. *Social Science Research*, 21, 84–112.
- DiPrete, T. A. and Eirich, G. M. (2006). Cumulative disadvantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*, 32, 271–297.
- Downey, D. (1995). When bigger is not better: Family size, parental resources, and children's wellbeing. *American Sociological Review*, 60, 746–761.
- Ermisch, J. and Francesconi, M. (2013). The effect of parental employment on child schooling. *Journal of Applied Econometrics*, 28, 796–822.
- Freedman, D. A., Pisani, R., and Purves, R. (2007). *Statistics*. New York: W.W. Norton & Company.
- Frisell, T., Öberg, S., Kuja-Halkola, R., and Sjölander, A. (2012). Sibling comparison designs: Bias from non-shared confounders and measurement error. *Epidemiology*, 23, 713–720.

- Guo, G. and VanWey, L. (1999). Sibship size and intellectual development: Is the relationship causal? *American Sociological Review*, 64, 207–231.
- Griliches, Z. (1979). Sibling models and data in economics: Beginning of a survey. *Journal of Political Economy*, 87, S37-S64.
- Hauser, R. M. and Mossel, P. A. (1985). Fraternal resemblance and intersibling effects in educational attainment. *American Journal of Sociology*, 91, 650–673.
- Hauser, R. M. and Wong, R. S.-K. (1989). Sibling resemblance and intersibling effects in educational attainment. *Sociology of Education*, 19, 149–171.
- Härkönen, J. (2014). Birth order effects on educational attainment and educational transitions in West Germany. *European Sociological Review*, 30, 166–79.
- Hillmert, S. and Jacob, M. (2010). Selections and social selectivity on the academic track: A life course analysis of educational attainment in Germany. *Research in Social Stratification and Mobility*, 28, 59–76.
- Hsin, A. (2012). Is biology destiny? Birth weight and differential parental treatment. *Demography*, 49, 1385–1404.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., and Michelson, S. (1972). *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Kalmijn, M. and Kraaykamp, G. (2005). Late or later? A sibling analysis of the effect of maternal age on children's schooling. *Social Science Research*, 34, 634–650.
- Kristensen, P. and Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, 316, 1717.
- Kuo, H.-H. D. and Hauser, R. M. (1995). Trends in family effects on the education of black and white brothers. *Sociology of Education*, 68, 136–160.
- Lareau, A. (2011). *Unequal Childhoods: Class, Race, and Family Life. Second Edition*. Oakland, CA: University of California Press.
- Mare, R. D. and Tzeng, M.-S. (1989). Fathers' ages and the social stratification of sons. *American Journal of Sociology*, 95, 108–131.
- Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21, 685–701.
- McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography*, 41, 607–627.
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26, 67–82.

- Nicoletti, C. and Rabe, B. (2013). Inequality in pupils' test scores: How much do family, sibling type and neighbourhood matter? *Economica*, 80, 197-218.
- Pettersson-Lidbom, P. and Skogman Thoursie, P. (2009). Does child spacing affect children's outcomes? Evidence from a Swedish reform. *Institute for Evaluation of Labour Market and Educational Policy Working Paper*, No. 2009-7.
- Powell, B. and Steelman, L. C. (1990). Beyond sibship size: Sibling density, sex composition, and educational outcomes. *Social Forces*, 69, 181–206.
- Powell, B. and Steelman, L. C. (1993). The educational benefits of being spaced out: Sibship density and educational progress. *American Sociological Review*, 58, 367–381.
- Powell, B. and Steelman, L. C. (1995). Feeling the pinch: Age spacing and economic investments in children. *Social Forces*, 73, 1465–1486.
- Powell, B., Steelman, L. C., and Carini, R. M. (2006). Advancing age, advantaged youth: Parental age and the transmission of resources to children. *Social Forces*, 84, 1359–1390.
- Rabash, J., Leckie, G., Pillinger, R. and Jenkins, J. (2010). Children's educational progress: Partitioning family, school and area effects. *Journal of the Royal Statistical Society, Series A*, 172, 657–82.
- Restrepo, B. 2012. Who compensates and who reinforces? Parental investment responses to child endowment shocks. *Unpublished Manuscript*, available at [http://www.mwpweb.eu/1/154/resources/research\\_255\\_1.pdf](http://www.mwpweb.eu/1/154/resources/research_255_1.pdf), accessed August 26, 2015.
- Schnitzlein, D. D. (2014). How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark. *Journal of Population Economics*, 27, 69–89.
- Schupp, J. and Herrmann, S. (2009). Kognitionspotentiale Jugendlicher: Ergänzung zum Jugendfragebogen der Längsschnittstudie Sozio-oekonomisches Panel (SOEP). *SOEP Data Documentation*, No. 43.
- Sieben, I., Huinink, J., and de Graaf, P. (2001). Family background and sibling resemblance in educational attainment: Trends in the former FRG, the former GDR and the Netherlands. *European Sociological Review*, 17, 401–430.
- Solon, G., Corcoran, M., Gordon, R., and Laren, D. (1991). A longitudinal analysis of sibling correlations in economic status. *Journal of Human Resources*, 26, 509–534.

- Solon, G., Page, M. E. and Duncan, G. J. (2000). Correlations between neighboring children in their subsequent educational attainment. *Review of Economics and Statistics*, 82, 383-392.
- Teachman, J. D. 1995. Sibling resemblance and symmetry in intellectual skill. *Sociology of Education*, 68, 205–220.
- TNS Infratest Sozialforschung (2013). SOEP 2012 – Methodenbericht zum Befragungsjahr 2012 (Welle 29) des Sozio-oekonomischen Panels. *SOEP Survey Papers*, No. 144.
- Toka, G. and Dronkers, J. (1996). Sibling resemblance in educational attainment, occupational prestige, and wealth in Hungary during the communist regime. *European Sociological Review*, 12, 251–269.
- Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40, 1008–1018.
- van Eijck, K. and de Graaf, P. M. (1992). The effects of family structure on the educational attainment of siblings in Hungary. *European Sociological Review*, 11, 273–292.
- Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP): scope, evolution and enhancements. *Schmollers Jahrbuch (Journal of Applied Social Sciences)*, 127, 139–169.
- Warren, J. R., Sheridan, J. T., and Hauser, R. M. (2002). Occupational stratification across the life course: Evidence from the Wisconsin Longitudinal Study. *American Sociological Review*, 67, 432–455.



# CHAPTER 6: DISCUSSION AND CONCLUSION

## Summary of Results

Do the effects of disadvantageous family and sibling characteristics vary by family socio-economic background? This thesis has analyzed the social origin heterogeneity in the effects of month of birth, parental separation, maternal age, close birth spacing, and birth order on child education.

Table 6.1 provides an overview over the main findings of the four empirical studies collected in this thesis with respect to the heterogeneity in these effects. Apart from reporting the main findings from the four empirical chapters, this table also reports the key identifying assumptions which underlie the research designs employed in the different chapters.

**Table 6.1** Main findings of the thesis and identifying assumptions of research designs

Chapter	Explanatory variable(s)	Main finding(s)	Key identifying a ssumption
2	Month of birth	Negative effect of school entry age on school performance stronger in families with a low level of parental education	Independence of month of birth
3	Parental separation	Negative effect of parental separation on track attendance and school grades only in families with a low level of parental education	No unobserved sibling-specific influences
4	Parental separation	Negative effect of parental separation on school grades only in families with a low level of parental education	Independence of ratio of opposite sex colleagues at the maternal workplace
5	Birth order, close birth spacing, maternal age	Negative effect of close birth spacing on track attendance only in families with a low level of parental education and a low parental class	Further parental fertility progression not influenced by child outcomes

The first empirical investigation, reported in Chapter 2, has used data on England. It followed the effects of month of birth, which determines the school entry age, throughout the school career of English pupils. It has shown that there is no gap between lower and higher social origin families in the effects of school entry age before school start but an increasing gap during the school career.

The second empirical study, reported in Chapter 3, has applied family-fixed effects models to the study of the effects of parental separation on child education using data on

Germany. The results of this study suggest that educational outcomes are only negatively affected by parental separation for children from families with a low level of parental education.

The third empirical study, reported in Chapter 4, uses a new IV approach to study the effects of parental separation on child education in Sweden. Its results, very much in line with the results of the previous chapter suggest that parental separation negatively influences child education only in families with a low level of parental education.

The fourth empirical study, reported in Chapter 5, looks at the role of within- and between-family inequality in educational outcomes in Germany. This study finds that the role of within- and between-family inequality is similar in lower and higher social origin families. However, the negative effects of close birth spacing on educational outcomes are only brought about in low-SES families.

### **Theoretical Implications for Research on Educational Inequalities**

This thesis is a collection of four empirical studies; these studies make individual contributions to research on the effects of family and sibling characteristics on educational outcomes and the variation of these effects with social origin. The contribution each chapter makes to the research question analyzed in a specific chapter is discussed in the conclusion of each empirical chapter. However, there are some common points connecting the four studies which I would like to point out.

First, the papers share a common theoretical framework which guided the analysis presented in all four empirical applications. The theoretical framework tested in this thesis was the compensatory effect of social origin hypothesis. I could confirm this framework for the effects of school entry age, parental separation, and close birth spacing on child education. No support for the hypothesis was found with respect to the effects of birth order and maternal age. These examples show that a compensatory effect is not observed with respect to every family characteristic which influences educational outcomes.

However, I believe that my results demonstrate that the compensatory effect is a frequent phenomenon which deserves further attention and explanation. Also because other research demonstrates a similar result with respect to other disadvantageous characteristics or life events (e.g. Almond et al. 2009; Bernardi 2014; Ermisch and Francesconi 2013; Torche and Echevarría 2011).

The exceptions to the compensatory effect of social origin hypothesis are interesting topics for further research. A general theory of social origin differences in education should explain in which instances the compensatory effect emerges and in which it does not. But the exceptions to this theory should also allow researchers to shed light on the underlying mechanisms of the compensatory effect. With respect to the results of this paper the main exceptions to the compensatory effect of social origin hypothesis are birth order and maternal age effects. Further research may also explain why the compensatory effect occurs with respect to some family characteristics but not to all of them.

Second, the results of this thesis have important implications for research on the intergenerational transmission of education (Breen and Jonsson 2005). I argued in the introduction to this thesis that the differential reactions by different social classes to disadvantageous life events provide an additional mechanism which leads to the high inequalities in educational outcomes by social origin in contemporary societies. The results of this thesis suggest that research which tries to explain the intergenerational transmission of education has to take into account the influence of the family and sibling characteristics analyzed in this study.

Third, the results of the studies collected in this thesis contribute to theoretical debates about the effects of the family and sibling characteristics under study. With respect to the “diverging destinies” (McLanahan 2004) hypothesis, the results suggest that not only disadvantageous life events are more common to occur in lower class families. The consequences of experiencing them are also more negative for children from lower class families with respect to their educational outcomes. For instance, the results on separation effects suggest that these effects differ substantially between families with few and families with many resources.

Fourth, my results on differences between siblings contribute to the further development of theories on intra-family resource allocation. The results with respect to sibling correlations go in slightly different directions than results for sibling correlations in the US (Conley and Glauber 2008; Conley et al. 2007). This may point to underlying cross-national differences which could be investigated by further research.

### **Methodological Implications for Research on Educational Inequalities**

The empirical analyses reported in this thesis share the methodological conviction that in order to isolate the effects of family characteristics on educational outcomes it is important to

control for the confounding influence of unobserved variables. To my mind, research on social stratification can gain a deeper understanding of the underlying causal mechanisms through the use of these techniques. Such research designs allow researchers to interpret estimates as causal effects under weaker assumptions, which is an important aim of social science research.

In the individual chapters I have shown that conclusions based on research designs which control for the influence of unobserved variables that cannot be measured by the researcher can differ from conclusions based on the naive estimates. This thesis, hence, questions the use of such naive estimation methods to study the effects of family and sibling characteristics on child education and the compensatory effect of social origin. If at all, I would argue that further research should move in the direction of using research designs that allow researchers to interpret causal effects under weaker assumptions than the designs employed in this thesis. For instance, the family-fixed effects models applied in Chapter 2 and in Chapter 4 rely on the assumption that parents do not take the abilities of their children into account when they make decisions on their further fertility progression. This is a rather strong assumption which could be tested by further research.

In any case, a return to the naive estimation strategies does not improve on the analytical strategies employed in this thesis. Not only do research designs which do not control for selection on unobserved variables estimate biased effects, the direction of the bias is also unpredictable. In my view, this makes these methods inappropriate for causal analysis and their use should remain restricted to supporting descriptive claims.

Although research designs which control for unobserved heterogeneity can show that selection on unobserved variables into parental separation takes place, they do not allow researchers to determine which these unobserved variables are. Only in very lucky circumstances the specific variables leading to selection could be identified. With respect to the applications in this thesis this is, however, not possible and it can only be speculated which variables lead to selection effects.

This point can be illustrated with the example of parental separation. Its effects are analyzed in two studies included in this thesis. Selection into parental separation may be driven by factors such as a weak interest of the parents in educating their children, which also directly affect child education. This selection may also differ between social origin groups. The results of the analyses reported in Chapter 3 and in Chapter 4 suggest that selection on unobserved variables is mainly happening in families with a high level of parental education. This is why no negative effect of parental separation is found for children from these families.

This differential selection into parental separation demonstrates again that the compensatory effect of social origin cannot be studied without a research design which ensures that the characteristic under study is exogenous to social origin.

But such considerations about the mechanisms leading to selection remain subject to speculation unless tested themselves. Such a test is, however, not necessary to understand the results. The power of causal identification strategies lies precisely in the fact that they control for the influence of unobserved, confounding variables without the need of knowing how they operate.

### **Mechanisms Underlying the Compensatory Effect of Social Origin**

This thesis has demonstrated that the consequences of some disadvantageous life events and characteristics are more negative for children from lower class than for children from higher class families. This section discusses the possible mechanisms underlying the heterogeneity in these effects and the circumstances under which further research could test them.

Parental involvement may be a likely mechanism connecting the effects of family and sibling characteristics, their variation by social origin, and educational outcomes. Research has persistently shown a positive relation between measures of parental involvement and children's educational outcomes (Chan and Koo 2011; El Nokali et al. 2010; Ermisch 2008; Harris et al. 1998; Kalil 2014; Kaushal et al. 2011; Yeung et al. 2002; Wang et al. 2014). In addition, parental involvement varies with social origin in a way that those types of parental involvement which are supposedly positive for children's outcomes occur more often in upper than in lower class families (Bianchi et al. 2004; Gracia 2014; Guryan et al. 2008; Hoff 2003; Hsin and Felfe 2014; Kalil et al. 2012; Lareau 2011). These two robust findings of previous research make parental involvement a likely mechanism that may underlie the compensatory effect of social origin.

In principle, the compensatory effect could be a consequence of differences in parental involvement for two reasons. Responses within lower and upper class families vary in a way which could explain this finding, either through a compensatory response by socio-economic advantaged families or through a reduction of parental involvement in lower class families in response to a disadvantageous life event.

Nevertheless, it should also be acknowledge that there is a possibility that the observed compensatory effect is due to a mechanical response without that parental involvement is involved (Almond et al. 2009). It may be that low-SES children just have a lower level of

abilities with a disadvantageous life pushing them below a threshold. Children from upper class families have, however, a higher level of abilities making it less likely that they are pushed below a threshold by a disadvantageous life event or characteristic. Since this possibility exists, it is important for theoretical explanations of educational inequality and for policy recommendations alike, to test empirically whether parental involvement responses are indeed causing the compensatory effect.

With respect to parental involvement as one mechanism underlying the compensatory effect of social origin, two studies on the effect of maternal employment on child outcomes are of interest. Ermisch and Francesconi (2013) show that the negative effect of maternal employment on child education is stronger in low social origin families. This result is fully in line with the compensatory effect of social origin hypothesis. Moreover, Hsin and Felfe (2014) show that children in low social origin families spend less educational time with their mothers as a consequence of maternal employment and that educational time positively affects child cognitive development. In combination, the studies suggest that parental involvement may transmit the compensatory effect of social origin, although neither of these studies actually shows this by including both the effect of maternal employment on child education and its variation with social origin as well as the effect of maternal employment on parental involvement and its variation with social origin in the same analysis.

Similar results are reported with respect to the influence of birth weight on children's educational outcomes. Torche and Echevarría (2011) show that the negative effect of a low birth weight on child cognitive outcomes is concentrated in families with low educated mothers. Both Hsin (2012) and Restrepo (2012) show that the negative effect of a low birth weight on parental involvement, measured through maternal time spent with the child (Hsin 2012) and through the HOME observatory (Restrepo 2012), are stronger in families with low educated parents. And again in this case, neither of these studies shows that the social origin differential in the effect of birth weight on parental involvement actually explains the social origin differential in the effect of birth weight on child education.

In Chapter 2 of this thesis I could not produce any evidence that parental involvement is underlying the social origin differential in the compensatory effect with respect to school entry age. We tested whether parental help with homework and paying for private lessons mediates the compensatory effect. However, we did not find any evidence for this mediation. Maybe there is no such mediation or we may not have used the appropriate measure to test the mediation. Further research will have to deal with that question and may, hopefully, employ better measures of parental involvement. Our test reported in Chapter 2 assumes that behavior

of high-SES families is driving the compensatory effect of social origin. It could, however, also be behavior of low-SES families that drives the observed compensatory effect.

With respect to parental separation, I find evidence for a social origin differential in parental involvement following parental separation in a recent paper. Using the same research design and data as in Chapter 3 (family-fixed effects models and data from the German Socio-Economic Panel Study) I show that parental separation only affects father involvement in families with a low level of parental education (Grätz 2015). In this study, I am also not able to test whether these changes in paternal involvement following parental separation explain the social origin differential in the effect of parental separation on child education since parental involvement and educational outcomes are measured at the same time.

Moreover, there are threats to the identification of causal mechanisms. A classic way to test the mediating role of mechanisms has been mediation analysis (Baron and Kenny 1986; Blau and Duncan 1967). However, the application of mediation analysis within a design-based approach framework, such as it was applied throughout most of this thesis, is far from being straightforward. Mediation analysis is based on additional assumptions in order to allow researchers a causal interpretation (Bullock et al. 2010; Imai et al. 2011; MacKinnon and Pirlott 2015).

Let us consider again the example of the analysis of the influencing role of private lessons and parental help with homework discussed in Chapter 2 of this thesis. The logic which is followed in testing the influencing role of these two potential mediating variables is by adding them to the regression models as additional control variables. The underlying idea of this approach is the argument that, in case the interaction effect between relative age and parental education is reduced once these variables are added to the model, these variables are underlying, at least partly, the compensatory effect of social origin.

It should, however, be noted that such an interpretation relies on additional assumptions. Such an analysis has to show that month of birth influences child education and that month of birth leads to changes in parental involvement. Furthermore, interpreting the mediating effect still assumes that no other variable, which would also constitute a plausible mechanism, changes in the same way as the analyzed mechanisms. This assumption has been labeled the sequential ignorability assumption by Imai et al. (2011).

Causal identification strategies may then be precisely used to study the “causes of effects” (Goldthorpe in press) or “forward causal questions” (Gelman and Imbens 2013). Goldthorpe (in press) acknowledges that experiments are one way to test the underlying mechanisms. This demonstrates, once more, that design-based approaches to causal inference

and mechanisms as the generative processes assumed to bring about the causal effects are not mutually exclusive but complementary approaches, as this has also been argued in Morgan and Winship (2015).

Future work may hopefully prove this point not only in theory but also in practice through the identification of causal effects and the underlying causal mechanisms of these causal effects. Such a causal mediation analysis would have to show three relations. First, a causal effect of a life event on child education would have to be identified. Second, it would have to show that the life event causes a change in the mechanism (e.g. parental involvement) which is supposed to operate. Third, the analysis needs to demonstrate that the change in the mechanism, and not some other change that occurs parallel, brings indeed about the effect of the life event on child education (Imai et al. 2011).

The compensatory effect of social origin hypothesis argues that the working of these mechanisms differs between lower and upper class families and conducting the analysis separately on a lower and a higher class sample could help identify whether the compensatory effect is caused by the behavioral responses of the lower or of the upper class.

## **Limitations**

Before policy recommendations are discussed in the next section, I believe it is important to point out limitations in the analysis presented in this thesis. These limitations should be kept in mind not only so that the validity of the findings can be examined but also so that policy recommendations can be deducted from them. There are, in particular, three limitations which I would like to point out.

First, each analysis in this thesis relies on specific assumptions which allow me to interpret the obtained estimates as identifying underlying causal effects. The last column of Table 6.1 reports the key assumptions made in the specific analyses. It has to be decided with respect to each chapter separately whether these assumptions hold. To the best of my knowledge there is no evidence available yet which would contradict any of these assumptions. However, future research may invalidate some assumptions. In that sense the results of these papers are preliminary and are open to revision by research that can demonstrate that an assumption does not hold and that this leads to biased estimates.

Second, each chapter reports a specific study conducted with respect to one cohort in one country. Extrapolating results of a specific study to cohorts and countries that were not included in the analysis, should be avoided (Freedman 2009). It is impossible to rule out that

in other settings effects may turn out differently. To my mind, external validity cannot be achieved by a single study but only through common efforts of the scientific community. For this reason, I hope other researchers will take up the challenges provided by my analyses and try to replicate the findings using alternative causal identification strategies and data on different cohorts and countries.

Third, a further limitation of the chapters presented in this thesis is the neglect of cross-national variation. I decided to concentrate on single country studies which isolate the effects of family characteristics on educational outcomes. In the true world, however, contextual factors such as the educational system of a country may have an effect on how these family characteristics affect educational outcomes of children and the variation of these effects by family socio-economic background. Again, to my mind, this limitation can only be resolved if others reproduce results using data on different countries. There is a lot of research in sociology on the effects of educational systems and other macro-level factors. In my defense, however, it can be said that the causal effects these influencing factors have, are often unclear because of research designs which rely on strong assumptions that are not very likely to hold up.

An example of how to deal with the second and third limitation is provided in this thesis through the analyses reported in Chapter 3 and 4. In these chapters I present results on the effects of parental separation on child education using data from two different countries as well as different identification strategies. Nevertheless the results of both studies are very similar suggesting that the effects of parental separation do not differ for recent cohorts in Germany and in Sweden. Further research must test whether similar results can be obtained for other countries and if the effects of other influencing factors differ in other countries.

Coming back to the last limitation, I believe that there is a lot of potential for future research combining the dimension of macro-level explanatory variables with the identification of causal effects at the micro-level. However, even finding data on two countries which would allow researchers to implement a causal identification strategy is a very difficult task. But still I believe that it is a first important step if such research can show whether causal effects vary or do not vary between countries before we can turn to analyze whether institutions themselves have a causal effect in moderating these micro-level effects. This is an even bigger challenge since it requires identifying two causal effects simultaneously, one at the micro and one at the macro level (Almond and Mazumder 2013).

## **Policy Recommendations**

The primary aim of this thesis was to explore the influence of family characteristics on educational inequalities and their variation by social origin. Policy recommendations have not been the primary focus of this thesis and I would be cautious in proposing too detailed policy recommendations which are not supported by the results presented in the empirical chapters of this thesis.

However, I have shown that several disadvantageous characteristics and life events (school entry age, parental separation, and close birth spacing) have mainly negative consequences for educational outcomes in lower social origin families. In that case, a policy recommendation may be that not everyone needs support if confronted with such a disadvantageous life event. Instead policies may be conceived which specifically target those who experience the disadvantageous life event *and* come from socio-economic disadvantaged families. This would help to focus scarce resources on those who are affected in their long-term outcomes by the life event.

These policies could take various forms. For instance, with respect to the month of birth penalty more liberal admission policies could reduce the negative effect of school entry age. It has also been proposed to boost the test scores of younger children (Crawford et al. 2014). The results of this thesis suggests that such a boost would have to be stronger for children from lower social origin families since they are, in the long run, affected more negatively by a young school entry age.

The consequences of parental separation could be reduced by providing extra support to children who experience parental separation. Closely spaced siblings could be targeted simultaneously through educational policies. In all these instances, policy measures should at the same time lead to a reduction of educational inequalities by social origin since mainly children from low social origin families are negatively affected by these individual and family characteristics.

## **Ideas for Further Research**

To my mind there are at least four directions in which future research could develop the ideas, results, and conclusions of this thesis.

First, in my view, this thesis, along with the quoted research others have conducted, provides evidence that the compensatory effect of social origin is a frequent phenomenon

which contributes to educational inequalities and which deserves more attention. Further research may analyze more instances in which the compensatory effect emerges. It is interesting to find out for which life events and family characteristics a compensatory effect can (and cannot) be found.

Second, cross-national variations in the effects of family characteristics on child education and the variations of these effects by social origin could be explored. As said above, this would require to use the same identification strategy for several countries.

Third, further work could focus on testing which mechanisms underlie the compensatory effect of social origin. Parental involvement may be one plausible mechanism. In this respect, it is important, for both theoretical considerations and for policy recommendations, to find out whether the compensatory effect is a result of the behavior of high-SES or whether it is due to responses from low-SES families.

A fourth way in which the results of this thesis can be expanded, is through the contribution this thesis makes to theories on intra-family resource allocation, in particular in Chapter 5. There is a lot of empirical evidence showing that parental involvement varies between children from the same family (Datar et al. 2010; Del Bono et al. 2012; Frijters et al. 2013; Halla and Zweimüller 2014; Hsin 2012; Restrepo 2012). So far only little research tests empirically whether parental involvement differs by family socio-economic background and what the contribution of parental involvement is to the intergenerational transmission of education.

Competing theories of parental investment in children predict that parents either reinforce (Becker and Tomes 1976) or compensate ability differences between their children (Behrman et al. 1982). Conley (2004, 2008) introduces the notion that whether parents compensate or reinforce depends on family socio-economic background (see also Griliches 1979). Studies testing whether parents reinforce or compensate ability differences have produced mixed evidence. Some studies provide empirical support of reinforcing parental behavior (Datar et al. 2010; Frijters et al. 2013) and some of compensatory parental responses to ability differences between their children (Del Bono et al. 2012; Halla and Zweimüller 2014). In addition, Hsin (2012) and Restrepo (2012) provide evidence for a variation of parental behavioral responses to birth weight differences with social origin. According to their complementary results, lower social origin families reinforce and higher social origin families compensate a low birth weight.

Trying to generalize the findings of Hsin (2012) and Restrepo (2012) for birth weight to child ability differences Grätz and Torche (2015) apply twin-fixed effects models to data from

the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B). However, we find opposing results of parental involvement responses to differences in early abilities between twins in two respects. First, we find that parents in higher social origin families provide more support to the higher ability twin. In contrast, lower social origin families do not show a response to ability differences. Second, we do not find any evidence that parental involvement mediates the effect of early child ability on later cognitive performance.

It is not easy to bring together these diverging empirical findings in a general theory of parental behavioral responses. One possible explanation of these diverging findings would be provided by a theory according to which parental involvement responses do differ with respect to different family and sibling characteristics and life events (Almond and Mazumder 2013). For instance, parents could reply with compensation to birth weight differences between their children but could react with reinforcement to ability differences between them (Grätz and Torche 2015). These responses could be stratified by socio-economic status. Such a multi-dimensional approach to parental responses to their children's characteristics may be able to explain the diverging findings with respect to the compensatory effect of social origin hypothesis found in this thesis. There may be no stratification in parental responses to differences in the birth order between their children because parents may not perceive different birth orders of their children as being problematic. On the other hand, high-SES parents may respond to month of birth differences, in particular in a country in which the effects of school entry age are regularly discussed in the media and well-known by highly educated parents. Similar, following a parental separation high-SES parents may anticipate negative consequences and react to reduce them for their children.

To my mind, another issue of importance in this literature, which needs clarification, is the causal effect of parental involvement on children's educational outcomes. Previous research has mostly shown a positive association between measures of parental involvement and child education but has largely not employed research designs which allow researchers a causal interpretation under the same weak assumptions as the research designs that I used throughout this thesis (El Nokali et al. 2010; Harris et al. 1998; Kaushal et al. 2011; Milkie et al. 2015; Yeung et al. 2002; Wang et al. 2014). Three studies are exceptions in that respect. Ermisch (2008) employs an upper and lower bound estimation strategy which does not lead to precise point estimates about the effect size of the influence of parental involvement on child education. Hsin and Felfe (2014) find educational time spent with the mother causes higher child test scores in individual-fixed effects models. In our study which uses twin-fixed effects models, we find, however, no increase in child cognitive outcomes at age 4 years with more

parental involvement at age 2 years (Grätz and Torche 2015). The latter two studies differ from each other in several ways: they apply different causal identification strategies, the children have different ages when their outcomes are observed, and the studies measure parental behavior through different indicators. Further research may test which of these differences explains the differences in results.

One way to move forward with this research, in my view, is to combine the rich measures of parental involvement often employed in psychology with the rigorous causal analysis mainly employed in econometrics. The main insight sociology adds is the notion that these processes are likely to differ with social origin.

## REFERENCES

- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl's subclinical legacy: Prenatal exposure to radioactive fallout and school outcomes in Sweden. *Quarterly Journal of Economics*, 124, 1729–1772.
- Almond, D. and Mazumder, B. (2013). Fetal origins and parental responses. *Annual Review of Economics*, 5, 37–56.
- Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Becker, G. (1991). *A Treatise on the Family. Enlarged Edition*. Cambridge, MA: Harvard University Press.
- Becker, G. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of Political Economy*, 84, S143–S162.
- Behrman, J. R., Pollak, R. A., and Taubman, P. (1982). Parental preferences and provision for progeny. *Journal of Political Economy*, 90, 52–73.
- Bernardi, F. (2014). Compensatory advantage as a mechanism of educational inequality: A regression discontinuity based on month of birth. *Sociology of Education*, 87, 74–88.
- Bianchi, S., Cohen, P. N., Raley, S., and Nomaguchi, K. (2004). Inequality in parental investment in child-rearing: Expenditures, time, and health. In Kathryn M. Neckerman (Ed.), *Social Inequality* (pp. 189–220). New York, NY: Russell Sage Foundation.
- Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure*. New York, NY: Wiley.

- Breen, R. and Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, 31, 223–243.
- Bullock, J. G., Green, D. P., and Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, 98, 550–558.
- Chan, T. W. and Koo, A. (2011). Parenting style and youth outcomes in the UK. *European Sociological Review*, 27, 385–399.
- Conley, D. (2004). *The Pecking Order: Which Siblings Succeed and Why*. New York, NY: Pantheon Books.
- Conley, D. (2008). Bringing sibling differences in: Enlarging our understanding of transmission of advantage in families. In D. Conley and A. Lareau (Eds.), *Social Class: How Does It Work?* (pp. 179–200). New York, NY: Russell Sage Foundation.
- Conley, D. and Glauber, R. (2008). All in the family? Family composition, resources, and sibling similarity in socioeconomic status. *Research in Social Stratification and Mobility*, 26, 297–306.
- Conley, D., Pfeiffer, K. M., and Velez, M. (2007). Explaining sibling differences in achievement and behavioral outcomes: The importance of within- and between-family factors. *Social Science Research*, 36, 1087–1104.
- Crawford, C., Dearden, L., and Greaves, E. (2014). The drivers of month of birth differences in children's cognitive and non-cognitive skills: A regression discontinuity analysis. *Journal of the Royal Statistical Society: Series A*, DOI: 10.1111/rssa.12071.
- Datar, A., Kilburn, M. R., and Loughran, D. S. (2010). Endowments and parental investments in infancy and early childhood. *Demography*, 47, 145–162.
- Del Bono, E., Ermisch, J., and Francesconi, M. (2012). Intrafamily resource allocations: A dynamic structural model of birth weight. *Journal of Labor Economics*, 30, 657–706.
- El Nokali, N. E., Bachman, H. J., and Votruba-Drzal, E. (2010). Parental involvement and children's academic and social development in elementary school. *Child Development*, 81, 988–1005.
- Ermisch, J. (2008). Origins of social immobility and inequality: Parenting and early child development. *National Institute Economic Review*, 205, 62–71.
- Ermisch, J. and Francesconi, M. (2013). The effect of parental employment on child schooling. *Journal of Applied Econometrics*, 28, 796–822.
- Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge: Cambridge University Press.

- Frijters, P., Johnston, D. W., Shah, M., and Shields, M. A. (2013). Intra-household resource allocation: Do parents reduce or reinforce child ability gaps? *Demography*, 50, 2187–2208.
- Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105, 765–789.
- Gelman, A. and Imbens, G. (2013). Why ask why? Forward causal inference and reverse causal questions. *National Bureau of Economic Research Working Paper*, No. 19614.
- Goldthorpe, J. (in press). *Sociology as a Population Science*. Cambridge: Cambridge University Press.
- Gracia, P. (2014). Fathers' child care involvement and children's age in Spain: A time use study on differences by education and mothers' employment. *European Sociological Review*, 30, 137–150.
- Grätz, M. (2015). Leaving the partner, leaving the child? Are parents really less involved in their children's lives after they separate? *Unpublished Manuscript*.
- Grätz, M. and Torche, F. (2015). Compensating or reinforcing? Variation in parental involvement responses to early child abilities by social origin. *NYU Population Center Working Paper*.
- Griliches, Z. (1979). Sibling models and data in economics: Beginning of a survey. *Journal of Political Economy*, 87, S37–S64.
- Guryan, J., Hurst, E., and Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, 22, 23–46.
- Halla, M. and Zweimüller, M. (2014). Parental response to early human capital shocks: evidence from the Chernobyl accident. *IZA Discussion Paper*, No. 7968.
- Harris, K. M., Furstenberg, F. F., Jr., and Marmer, J. K. (1998). Paternal involvement with adolescents in intact families: The influence of fathers over the life course. *Demography*, 35, 201–216.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378.
- Hsin, A. (2012). Is biology destiny? Birth weight and differential parental investment. *Demography*, 49, 1385–1405.
- Hsin, A. and Felfe, C. (2014). When does time matter: Maternal employment, children's time with parents, and child development. *Demography*, 51, 1867–1894.

- Kalil, A. (2014). Inequality begins at home: The role of parenting in the diverging destinies of rich and poor children. In P. Amato, A. Booth, S. McHale, and J. Van Hook (Eds.), *Diverging Destinies: Families in an Era of Increasing Inequality* (pp. 63–82). New York: Springer.
- Kalil, A., Ryan, R., and Corey, M. (2012). Diverging destinies: Maternal education and the development gradient in time with children. *Demography*, 49, 1361–1383.
- Kaushal, N., Magnuson, K., and Waldfogel, J. (2011). How is family income related to investments in children's learning? In G. Duncan and R. Murnane (Eds.), *Whither opportunity? Social inequality and educational disadvantage* (pp. 187–206). New York, NY: Russell Sage Foundation.
- Lareau, A. (2011). *Unequal Childhoods: Class, Race, Family Life. Second Edition*. Berkeley, CA: University of California Press.
- MacKinnon, D. P. and Pirlott, A. G. (2015). Statistical approaches for enhancing causal interpretation of the M to Y relation in mediation analysis. *Personality and Social Psychology Review*, 19, 30–43.
- McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography*, 41, 607–627.
- Milkie, M. A., Nomaguchi, K. M., and Denny, K. E. (2015). Does the amount of time mothers spend with children or adolescents matter? *Journal of Marriage and Family*, 77, 355–372.
- Morgan, S. L. and Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research. Second Edition*. Cambridge: Cambridge University Press.
- Restrepo, B. (2012). Who compensates and who reinforces? Parental investment responses to child endowment shocks. *Unpublished Manuscript*, available at [http://www.mwpweb.eu/1/154/resources/research\\_255\\_1.pdf](http://www.mwpweb.eu/1/154/resources/research_255_1.pdf), accessed August 26, 2015.
- Torche, F. and Echevarría, G. (2011). The effect of birthweight on childhood cognitive development in a middle-income country. *International Journal of Epidemiology*, 40, 1008–1018.
- Wang, M.-T., Hill, N. E., and Hofkens, K. (2014). Parental involvement and African American and European American adolescents' academic, behavioral, and emotional development in secondary school. *Child Development*, 85, 2151–2168.

Yeung, J., Linver, M., and Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, 73, 1861–1879.