

Sibling Configuration and Adulthood Outcomes

The Case of Two-Child Families

Roujman Shahbazian

Academic dissertation for the Degree of Doctor of Philosophy in Sociology at Stockholm University to be publicly defended on Friday 1 June 2018 at 10.00 in Hörsal 11, Hus F, Universitetsvägen 10 F.

Abstract

This thesis includes three empirical studies, analyzing how sibling configuration (i.e. birth order, birth spacing and sex-composition) influences siblings' long-run income and educational choice. This is done by utilizing the unique linkage opportunities of administrative registers covering the entire population of Sweden.

Study I: This paper focuses on how different birth spacing intervals are associated with income rank from ages 33 to 42 years, for siblings in two-child families. The results show clear differences between first- and second-born siblings. At the more common spacing intervals (less than 5 years), spacing has a negligible association to second-born children's long-term income rank. However, first-born children have lower income rank when a younger sibling is born when they are very young. Having relatively high spacing intervals (over 5 years) is associated with somewhat lower long-term income-rank than having mid-length intervals for both first- and second-born siblings.

Study II: This study focuses on the association between combinations of sibling configuration (i.e. birth order, birth spacing and sex composition) and long-run income rank of siblings. The results show that the significance of different family factors in two-child families vary by sibling sex-composition. The findings suggest that both birth order and birth spacing are important factors for first born boys independent of the younger sibling's sex. First-born girls, however, only have an advantage if they have a younger sister. More surprisingly is that this advantage does not seem to vary by birth spacing.

Study III: This study examines how sibling gender configuration in Swedish two-child families influences the choice of so-called STEM educational fields (i.e. Science, Technology, Engineering and Mathematics). The results show that younger siblings, net of parental characteristics, are more likely to choose a STEM field if their older sibling already has attended a STEM program. The findings indicate that boys' choice of STEM fields is independent of having an older brother or sister who has attended a STEM program. However, girls seem to be more likely to choose a STEM-field if they have a sister who has attended a STEM program, than if they have a brother with a similar program. Given that STEM-fields are markedly male dominated, this indicate the importance of having a same-sex role model for making gender atypical educational choices.

Keywords: *educational choice, income, siblings, birth order, birth spacing, sex-composition, STEM, Sweden.*

Stockholm 2018

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-155155>

ISBN 978-91-7797-155-9

ISBN 978-91-7797-156-6

ISSN 0283-8222



Stockholm
University

Department of Sociology

Stockholm University, 106 91 Stockholm

SIBLING CONFIGURATION AND ADULTHOOD OUTCOMES

Roujman Shahbazian



Sibling Configuration and Adulthood Outcomes

The Case of Two-Child Families

Roujman Shahbazian

©Roujman Shahbazian, Stockholm University 2018

ISBN print 978-91-7797-155-9

ISBN PDF 978-91-7797-156-6

ISSN 0283-8222

Printed in Sweden by Universitetsservice US-AB, Stockholm 2018

Distributor: Swedish Institute for Social Research, Stockholm University

Cover photography: Yvonne Naderi

Contents

Acknowledgements	2
Abstract	4
Svensk sammanfattning	6
Introduction	7
Theoretical Background	8
Social learning	9
Confluence model and resource dilution model	10
Niching	11
Previous Studies on sibling configuration	11
Adulthood outcomes	13
Long-run income	13
Tertiary educational field	14
Methods and data: Strengths and limitations	15
Data and interaction models	15
Analytical strategies when studying sibling configuration	16
Overview of the studies	18
References	20

Acknowledgements

During the last years, I have accumulated many debts to those who have paved the way for this thesis, by directly or indirectly contributing to it.

I would like to start by thanking the Swedish taxpayers for paying their taxes each month, which financed my PhD position¹, and different government agencies, such as Statistics Sweden, for collecting, harmonizing and storing the administrative registers used in this thesis. Without their effort, the content of this thesis would not have been materialized.

I owe a great deal to the Level-of-Living (LNU) unite at SOFI, for accepting and educating a doctoral student who had no previous sociological training. The main responsibility for this task, has fallen on my supervisors, which I am deeply indebted to. Magnus Neremo has been a source of inspiration and encouragement. As a supervisor, he has been more than generous with his time, and provided me with plenty of freedom to pursue my own research interests. The first person I got to know at SOFI was Erik Bihagen, and I am lucky that he has shared his knowledge in various aspects of doing research with me. Katarina Boye has been a great support with her patience and ease in answering questions, both major and minor, from her self-doubting PhD student. In the beginning of my PhD I had Carina Mood as co-supervisor. I learnt a great deal from her ability to combine the big picture with meticulous attention to detail.

I am fortunate to have been in a dynamic and constructive research environment as SOFI, which consists of both sociologists and economists. SOFI has created a constructive and friendly intellectual atmosphere to present and discuss each other's work, which is beneficial for junior researchers. I am also happy to have spent some time at the department of sociology at Stockholm University. There I meet skilled sociologists and demographers who were welcoming and generous with their time and knowledge. A special thanks goes to all the participants of the PhD lunch workshop during my years at SOFI. Many of you have been great company both in and outside of the university. I also wish to express my appreciation to the administrative staff, both at SOFI and the Department of Sociology, for helping me with administrative matters.

¹ Financial support from Swedish Research Council (VR) (Dnr: 2008:7499), the Swedish Research Council for Health, Working Life and Welfare (FORTE) (dnr: 2013-1119, and dnr: 2006-1515), is gratefully acknowledged.

I would like to single out Kieron Barclay and Juho Härkönen for their helpful comments on earlier versions of the thesis, at my half-time and final seminars. To Linus Andersson, Eva-Lisa Palmtag and Sara Kjellsson – thanks for checking parts of the manuscript at the eleventh hour. All shortcomings are definitely mine.

To my partner in crime, Mohammad Sepahvand. I am grateful for having had access to your sharp intellect, your wit, and your ability to offset my grumpy mood during this journey. I am looking forward to our future adventures!

During the last years, I have to some degree neglected many of my near and dear friends/relatives, which I hope to be able to spend more time with. To Tommy – you provide the rum and I will provide the whiskey glasses, deal? To Shant – so many AC Milan games, so little time (OK, perhaps also ONE Liverpool game...). To Iman – when should we go to the kaleh pacheh place? To my “siblings” (Ronak and Said) – I miss your good food. To Sohrab – hope to hear more of your plans. To my aunt Mehrangiz – count on more Sunday teas! To my uncle Jahanbakhsh – looking forward to our discussions about current and historical events, which is one reason I chose the path I am on.

Last but not least, I am eternally grateful to my parents, *Fereshteh* and *Ghodrat*, who have been my *Guardian Angel* and *Strength* during difficult times.

Uppsala, April 2018
Roujman Shahbazian

Abstract

This thesis includes three empirical studies, analyzing how sibling configuration (i.e. birth order, birth spacing and sex-composition) influences siblings' long-run income and educational choice. This is done by utilizing the unique linkage opportunities of administrative registers covering the entire population of Sweden.

Study I: Birth spacing between siblings may have long-lasting impact on them. This paper focuses on how different birth spacing intervals are associated with income rank from ages 33 to 42 years. In order to disentangle birth spacing from birth order while holding potential sibsize association constant, an interaction model is used on a restrictive subpopulation of two child families born between years 1960 and 1970. The results show clear differences between first- and second-born siblings. At the more common spacing intervals (less than 5 years), spacing has a negligible association to second-born children's long-term income rank. However, first-born children have lower income rank when a younger sibling is born when they are very young. Having relatively high spacing intervals (over 5 years) is associated with somewhat lower long-term income-rank than having mid-length intervals for both first- and second-born siblings.

Study II: Using longitudinal Swedish population data, we study the association between combinations of factors in the family of origin (i.e. birth order, birth spacing and sex composition) and the long-run income of siblings, measured as income rank averages between 33 and 42 years of age. Using interaction models, we find that the significance of different family factors in two-child families vary by sibling sex composition. The findings suggest that both birth order and birth spacing are important factors for first born boys independent of the younger sibling's sex. First-born girls, however, only have an advantage if they have a younger sister. More surprisingly is that this advantage does not seem to vary by birth spacing.

Study III: This study examines how sibling gender configuration in Swedish two-child families influences the choice of so-called STEM educational fields (i.e. Science, Technology, Engineering and Mathematics). By using population data, we find that younger siblings, net of parental characteristics, are more likely to choose a STEM field if their older sibling already has attended a STEM program. We also find a gender difference

among younger siblings choice of STEM fields. The findings indicate that girls are more likely to choose a STEM-field if they have a sister who has attended a STEM program, than if they have a brother with a similar program. Given that STEM-fields are markedly male dominated, this indicate the importance of having a same-sex role model for making gender atypical educational choices.

Svensk sammanfattning

Denna avhandling består av tre fristående studier som alla undersöker sambandet mellan syskonkonfiguration (dvs. födelseordning, åldersskillnad och könssammansättning), individers ackumulerade inkomst och utbildningsval i tvåbarnsfamiljer i Sverige. Detta studeras med hjälp av svenska registerdata där vi kan länka individer till deras förälder och därmed följa syskon över tid för att studera utbildningsval och inkomst i vuxen ålder.

Artikel 1 fokuserar på hur åldersskillnaden mellan syskon, d.v.s. födelseintervallet, påverkar rankningen av individers ackumulerade inkomst (medelinkomsten mellan 33 och 42 år). Av resultaten framgår att det finns en inkomstskillnad mellan först- och andrafödda syskon i tvåbarnsfamiljer. Förstfödda syskon tycks ha en fördel som ökar med åldersskillnaden mellan dem och deras yngre syskon. Detta gäller dock bara till en åldersskillnad på cirka 3 år. Andrafödda syskons inkomstrank tycks dock inte påverkas nämnbart av åldersskillnaden till det äldre syskonet. Enda undantaget är att deras inkomstank tycks sjunka något i de fall då åldersskillnaden är större än fem år.

Artikel 2 utgår från resultaten i artikel 1. Frågan här är dock om könssammansättningen mellan syskonen spelar roll för deras inkomstrank. Resultaten indikerar att födelseordning och åldersskillnad är viktiga faktorer för förstfödda pojkar, oberoende av deras yngre syskons kön. Förstfödda flickor har däremot endast en fördel om de har en yngre syster, och denna fördel tycks inte variera med åldersskillnaden mellan syskonen.

Artikel 3 fokuserar på hur äldre syskon påverkar sina yngre syskon när det kommer till att välja så kallade *STEM*-utbildningar (*Science, Tecknology, Engineering & Mathematics*) vid universitet eller högskolor. Av studien framgår att individer vars äldre syskon läst en *STEM*-utbildning är mer benägna än andra att själva välja en liknande utbildning. Flickors val av *STEM*-utbildning påverkas mer av ha en äldre syster som läst en *STEM*-utbildning, jämfört med att ha en äldre bror som läst en *STEM*- utbildning.

Introduction

The importance of family of origin for understanding intergenerational transmission of life chances and equality of opportunity is central to social stratification research. Most studies in this line of research focuses on how parents' socioeconomic standing, such as education, income and social class, influences children's adulthood outcome. This thesis is related to this field but instead of parents the focus is on aspects of sibling configurations, i.e. sibship size (the number of children in the family), birth order (the order a child is born in their family), child spacing (the amount of time between births), and sex composition (the sex composition of the sibling group). Studying aspects of sibling configurations is not a new field. It dates back to Galton's (1874) study of first-borns, and previous research in this field have made important contributions especially in areas related to school performance and educational attainment (see e.g. Bayer, 1967; Blake, 1981; Blake 1986; Downey 1995; Kuo and Hauser 1997). However, despite this long tradition, most of the above research has focused on either birth order or sibship size. Research on the association between birth spacing and adulthood outcome, and also the association between combinations of sibling configuration factors and siblings adult outcomes are still rare (for some important exceptions see e.g. Powell and Steelman, 1993; Buckles and Munnich, 2012; and Barclay and Kolk, 2017).

There are a number of reasons why sociologists should pay more attention to these factors. First, birth spacing between siblings is a factor that can be controlled by parents, through the use of contraceptives, and to some degree by governments through family policy (Hoem, 1993). Sociologists have therefore, as noted above, foremost shown an interest in analyzing how sibship size and birth order influence adulthood outcomes such as occupation and income opportunities (e.g. Blau and Duncan, 1967; Blake 1981; Steelman and Powell 1989; Parcel and Menaghan 1994). It is however more surprising that so little empirical attention has been paid to the impact of birth spacing. Second, although parents are the main agents in the primary socialization of children, McHale et al. (2012) argue that siblings' role as potential agents of socialization has been relatively neglected in the sociological literature.² Siblings engage in play with one another, they

² In their search of abstracts of scientific articles published in psychology and sociology between the years of 1990 and 2011, 741 citations were found for "sibling and relation or

confide and share secrets, fight with one another etc. Thereby, socialization takes place mostly as older siblings act as custodian of norms and values in the absence of parents. Even so, sociologists have with few exceptions paid little attention to study how different sibling configuration factors, such as gender and birth order, influence siblings' educational choices.³

This thesis therefore contributes to the field in three ways. First, by focusing on two-child families where siblings grew up together, we not only analyze the most common sibling group in Sweden, but also facilitate the theoretical and empirical understanding on how siblings interact with each other. Second, it fills a gap in the literature by not only studying the association between birth spacing and income rank in two-child families, but also the combined effect of several sibling configuration factors, i.e. birth order, birth spacing, and sex-composition on income rank. Third, the thesis also provides empirical evidence for the importance of older siblings for their younger sibling's educational choices, more specifically STEM programs (i.e. Science, Technology, Engineering and Mathematics), net of parental characteristics.

In the remainder of this introductory chapter, first the theoretical background is explained. Second, the previous research on sibling configuration and adulthood outcomes is reviewed. Third, the outcome measures in the thesis are presented and arguments are offered for why and how they are operationalized. Fourth, I discuss some strengths and limitations of using administrative data and methodological issues when focusing on two-child families. The introduction then ends with an overview of the three papers.

Theoretical Background

A common understanding of socialization is that it is being shaped by older members of a group: norms and values are transmitted from the older generation to the younger generation. This is also highlighted by the intergenerational literature in sociology where the family of origin plays a central role. Thus, parents can shape and transmit their preferences and attitudes to their children in various ways. Taking education as an example, in the home environment parents can guide children to set goals, encourage them to achieve them, monitor their progress and help them when needed (e.g., Erikson and Jonsson, 1996). Although the main agents of socialization in the family are the parents, siblings can also be agents of socialization for their siblings, as they tend to spend a large amount of time with each other (McHale et al. 2012). Thus, the fact that siblings tend to be similar to each

relationships”, in contrast to 33 990 citations for “parent or parenting” and 5 059 citations for “peer relations or peer relationships or friend-ships”.

³ For one important exception see Gabay-Egozi et al. (2016).

other in terms of e.g. educational attainment, income, and family formation (e.g. Björklund et al., 2009; Conley and Glauber, 2008; Raab et al., 2014), may also be an effect of sibling socialization. Understanding how sibling configuration in this sense facilitate siblings to act as agents of socialization, and also creates social environments where socialization can take place, can be approached in many different ways. Hence, sibling studies constitute a truly multidisciplinary field. The following section provides a brief description of the, in my view, major theoretical explanations to sibling configuration of high relevance in two-child families: i) social learning, ii) the confluence model and the resource dilution model, and iii) niching.

Social learning

One way to understand socialization is through asking who the socializing agents or the role models are. The short answer is that they can be individuals near the child, such as parents and neighbors, but also siblings. In families where siblings grow up together, older siblings are due to structural factors, such as birth order and age difference, likely to serve as a role model for their younger siblings.

According to social learning theory (Bandura 1977), later developed into social cognitive theory (Bandura, 1986), there are two processes by which children are influenced from role models. The first is instrumental training, where the role model is explicit in communicating what they wish the child to learn. This could be done through incentivizing children to acquire certain beliefs, attitudes and behavior. Depending on the characteristic of the role model, he or she possesses the power to reward certain behaviors and thereby their reproduction. The second is through a cognitive process where the child observes and imitates the social behavior of the role model. Role models are in this sense a communicative channel, i.e. if their actions are proven successful or performed frequently children are more inclined to imitate them. Imitating and following in the footsteps of role models does not need to be gender typical, but there are several elements which make it more likely for boys and girls to reproduce a gender appropriate behavior (Bussey and Bandura 1999). Thus, during the life-course individuals routinely experience certain type of activities performed by men, while other activities are performed by women. Taking division of housework in Sweden as an example, the time spent on gender typical activities is apparent not only among parents, but also among their children (Evertsson, 2006). This highlights not only the significance of imitation, but also the reproduction of gender typical behavior in the family and society. If individuals' educational and occupational choices are in this way limited by gender it generates a transition bias to positions in the society. Thus, individuals that are independent of gender more fit for a career would be hindered to aspire and ultimately attain it. From an individual level, systematic difference in how

women and men are treated may lead to unequal access to resources and opportunities, which in turn leads to gender differences in life-chances.

Confluence model and resource dilution model

In the introduction we noted that birth spacing between siblings may be important for understanding adulthood outcomes. The idea behind this is the confluence model, introduced by Zajonc and Markus (1975) and further developed by Zajonc (1976), which postulates that birth spacing is an important factor in the family structure to explain cognitive development. The age difference between siblings not only impacts the resources provided to children from parents but also the relationship between the siblings themselves. This could be illustrated by a family with one child where the child enjoys a rich intellectual environment, since the parents adjust their daily activities and conversation to the only child's needs. The arrival of a second child reduces the intellectual environment of the first child since the parental resources have to be divided between two children. However, depending on the age difference between the siblings, the effect of birth order might be mediated and the negative effect of family size can be compensated for, because the first-born child can benefit from tutoring the younger sibling, while the younger sibling benefits from having another individual (besides the parents) to learn from. According to the confluence model it can therefore be hypothesized that longer birth spacing between siblings are more advantageous for the older siblings cognitive development, but also the younger benefits since he or she has one extra "adult" to learn from (Zajonc 1976).

Another influential theory, the Resource dilution model, offers a related explanation for a potential difference in outcome between siblings due to birth spacing (e.g., Blake, 1981; Powell and Steelman, 1993; Downey, 1995). The main difference is that the resource dilution model does not argue that the primary channel for differences between siblings is the cognitive development of children. Instead, the primary channel has to do with the economical investment in children. Thus, according to the resource dilution model a short spacing between the siblings constrains the parental investment in children, especially in connection to more expensive investments such as tuition fees. Therefore, Kidwell (1981) argue that a wider spacing between siblings allows parents a *breathing room* when it comes to investment in children. However, since the Swedish society is characterized by a welfare state with heavily subsidized education and healthcare it is more likely that parental investment's as a mechanism proposed by resource dilution is less relevant than the cognitive development of children proposed by the confluence model.

Niching

The above theoretical explanations assume that the older sibling shape the attitudes and behaviors of the younger sibling to follow in their footsteps. However, the older sibling can of course also influence the younger to niche away, and especially when they are close in age. For instance, Sulloway (1996) propose that siblings consciously differentiate from each other, and that one reason behind this is that parental resources (i.e. time, attention, money etc.) are limited (or at least fixed), and that siblings therefore niche away in order to ease the competition between them. Another reason may be that younger sibling's niche away to find their own identity. Thus, niching can in this sense be regarded as a reaction against a sibling but also highly dependent on parents' actions and reactions, since one reason for niching is to gain parental resources. It is, in other words, reasonable to assume that the first-born sibling have due to their seniority an advantage when it comes to almost all aspects (such as hobbies etc.). If the second-born do not niche and instead engage in similar activities as their older sibling they might risk becoming "second", and in turn receive less appreciation by the parents. Thus, by niching, siblings avoid competing with each other and are more likely to be judged as equally "good" by their parents, and consequently receive the same amount of parental attention.

The niching hypothesis thereby contrasts the above explanations where the interaction between short birth spaced siblings is expected to make them more similar (see e.g. Zajonc, 1974).

Previous Studies on sibling configuration

Studies on different features of sibling configuration have, as mentioned above, a long history. Already, Galton (1874) made predictions about birth order and achievement based on his observations that members of the Royal Society in UK were disproportional first-born children. Over time other features of sibling configuration, such as sibship size, birth spacing and sex-composition, have also attracted researchers' attention (e.g. Blake, 1981; Powell and Stellman 1989; Powell and Stellman 1993; Downey, 1995). The following section briefly present more recent major empirical contribution to the field, related to sibship size, birth order, birth spacing, and sex composition.

One pattern repeatedly observed in the literature is that as the number of sibling's increases, younger siblings academic achievement decrease, in terms of e.g. educational attainment and/or grades (Blake 1981, Blake, 1986, Downey 1995, Kuo and Hauser, 1997). It has been argued that this effect of sibship size is endogenously determined by factors likely to influence siblings' educational outcomes, such as parental characteristics. Recent evidence have, however, indicated that the causal relationship between

sibship size and educational outcomes is either small (e.g., Black et al., 2005) or non-existing except for very large sibship sizes (e.g., Åslund and Grönqvist, 2010). However, since the overwhelming majority of families in Sweden are either one- or two-child families, sibship size is not a salient feature of sibling configuration in Sweden. Also, as noted above, the present thesis focuses on two-child families and thereby controls for the effect of sibship size. This is further discussed in data and methods section below.

The second sibling configuration factor discussed here is *birth order* and its association with achievement. The pattern that first born siblings are more successful than their younger siblings, when it comes to socioeconomic outcomes has been documented by many scholars (for an overview see Barclay, 2014). Even so, the strength and robustness of this birth order effect has been debated amongst scholars. Critics have argued that the association between birth order and educational outcomes is a methodological illusion either caused by poor data (e.g., Rogers et al. 2000), or insufficient controls for confounding factors (e.g., Hauser and Sewell, 1985; Rodgers, 2001; Steelman et al., 2002). However, more recent studies with access to better data on sibling groups combined with more sophisticated statistical methods indicate that there is a birth order effect (e.g. Black et al., 2005; Kantarevic and Mechoulan, 2006; Härkönen, 2013).

The importance of the third factor, *birth spacing*, has also for a long time attracted scholarly interest (e.g., Ansbacher and Rowena 1956; Zajonc and Markus 1975; Zajonc 1976), but it has until recently been less studied than both sibship size and birth order. However, Pettersson-Lindbom and Skogman-Thoursie (2009) found, by utilizing the Swedish parental leave benefit reform implemented in 1980, that close spacing between siblings (less than 2 years) had a negative effect on the older sibling's educational attainment. Moreover, Buckles and Munnich (2012) used the randomness of miscarriages as an exogenous variation in the length of spacing between siblings, and found that the test score of the older sibling increased for larger spaced siblings, defined as more than two years apart. Other researchers also report a positive association between birth spacing and educational achievement (Powell and Steelman 1993). However, by using sibling fixed effects, and focusing on families with three or more children in Sweden, Barclay and Kolk (2017) found no association between birth spacing and different socio-economic outcomes (measured between 29 and 31 years of age). Given this inconsistency in the literature, it is unclear whether birth spacing has an effect on adulthood outcomes. It could, however, be the case that the birth spacing effect wears off in larger family sizes. This assumption may then explain why studies report a birth spacing effect in two child families (e.g. Pettersson-Lindbom and Skogman-Thoursie, 2009; Buckles and Munnich, 2012), while birth spacing at the same time seem to have no effect for sibling groups of three or more (e.g., Barclay and Kolk, 2017).

Findings from studies focusing on the fourth and final factor, *sibling sex-composition*, are also mixed and even contradictory. Previous research report

a negative association for girls between having brother(s) and college enrolment (e.g. Powell and Steelman, 1989), or that girls with only brothers have more years of education compared to those who have at least one sister (Butcher and Case, 1994). Some studies indicate that the increased number of same-sex siblings is beneficial for both girls and boys educational attainment (Conley, 2000), while other studies find no association between sex-composition and educational outcomes (e.g. Hauser and Kuo, 1998; Amin, 2009) or income rank (Shahbazian, 2017). The majority of the above mentioned studies are based on survey data and as a result they run into the problem of analyzing cells with relatively few observations. Thus, it is difficult based on these studies to evaluate whether sex-composition in fact is associated with stratification outcomes.

In all, few studies combine several sibling configuration factors. Instead most studies, like the above, have tried to isolate one sibling configuration factor in order to control for all unobservable characteristics that may influence siblings' adulthood outcomes (e.g., Powell and Steelman, 1989; Butcher and Case, 1994; Conley, 2000; Buckles and Munnich, 2012). An alternative approach suggested in this thesis to overcome this problem is to use large datasets allowing for controls for potential confounding factors, e.g. parental characteristics, as well as being able to combine the above discussed sibling configuration factors. Thus, by using administrative registers, restricting the analytical population, and controlling for confounding parental characteristics, the internal validity of the results could increase.

Adulthood outcomes

Sibling configuration factors may be related to a number of adulthood outcomes. The papers included in this thesis focuses on two important outcomes that have long term consequences for individuals' well-being, i.e. long-run income, and tertiary educational choices.

Long-run income

Long-run income, used as outcome in Study I and II, can be viewed as the ultimate proxy of individuals' economic performance in the labor market, and is also a good indicator of inequality. The income measure used in the analyses is ranked, which enables us to study the association net of, e.g., structural changes over time, inflation and cohort effects. Income rank is also a good measure of the individual's specific position within the social hierarchy (cf. Carlsson et al. 2007), and therefore also a relevant measure for comparisons of siblings relative social standing. More specifically, the income rank is measured as each sibling's earned income rank between the

ages of 33-42 years of age, since previous research has shown that income for one year or at younger ages is a poor proxy for life-time income (Haider & Solon, 2006; Böhlmark and Lindquist, 2006). In order to be included in the analyses each sibling has to have at least 6 years of valid earned income in the registers.⁴ Earned income includes all taxable incomes from employment, self-employment, pension, sickness benefits and other taxable transfers.⁵ In a second step the income information was transformed into percentile ranks (1-100) based on the individuals earned income compared to all other individuals of the same age and sex in the Swedish labor market in a specific year. Thereafter, the dependent variable is defined by the average of the individual's income ranks (between ages 33-42). To give an example, Anna was 33 years old in 1995 and her income rank for that year is calculated by comparing her earnings to the earnings of all women in Sweden age 33 in 1995. Thus we do not wish to compare e.g. a 33-year-old to a 50-year-old, due to potential differences in the length of work experience. The same procedure has then been repeated until Anna was 42 years old. In a final step the average of these income ranks is calculated and used as a measure of her income rank between 33 and 42 years of age. Also, it is important to note that the above income ranks are calculated separately for women and men, since we do not wish income rank differences between siblings to be influenced by the existing gender wage gap.

Tertiary educational field

The second outcome variable used in this thesis, Study III, is a measure of the two sibling's tertiary educational choices. More specifically, it measures whether each sibling have chosen a tertiary education in a so called STEM field, i.e. a degree in Science, Technology, Engineering, or Mathematics. The reason for focusing on STEM fields is that they are highly demanded in the labor market and also argued to constitute the engine for a knowledge-based economy. Many countries and supranational organizations, such as the European Union, have therefore during the last decades put much emphasis and resources to increase the number of STEM majors (Caprile et al., 2015).

A second reason is that STEM fields include e.g. engineering and technical occupations which are not only markedly male dominated (Nermo

⁴ The reason for having a cut-off of at least 6 years of valid income information in the registers is that individuals might have been studying or working abroad during ages 33-42. However, over 90 percent of the sample has more than 9 years of valid income information. Sensitivity test have been performed to see whether there is an overrepresentation of older or younger siblings in the number of valid income information, the distribution is fairly equal. Sensitivity test have also been performed including siblings with 5 and 7 years of valid income information, without altering the result of the studies.

⁵ The variable does not include income from capital and non-taxable transfers. For more information about the income variable in Sweden from year 1968 and forward, see Hjalmarsson et al. (2015).

1999; Charles and Bradley, 2009; Mann and DiPrete 2013), but the return in terms of earnings are also better than for most other fields (Webber 2014). Thus, studying sibling influence in choosing STEM fields are thereby also interesting for understanding mechanisms related to gender atypical educational choices.

The variable measures whether the individual has been enrolled in a STEM program and has finished at least one semester at tertiary level.⁶ The educational field is collected between ages 33 and 42, depending on their year of birth.⁷ Educational attainment is here classified according to the Swedish national educational classification SUN⁸.

Methods and data: Strengths and limitations

This section presents a short discussion about the strength and limitation of administrative data, and how the analytical population has been restricted in order to adjust for potential confounding factors.

Data and interaction models

Administrative registers have many advantages compared to survey data. First, uncertainties when drawing inference is eliminated, as it covers the entire population and not only a random sample. Second, unlike survey data, register data of the entire population also allows for more complex models without running out of statistical power. Third, measures such as income and educational information, which to a large extent tend to be self-reported in survey data and thereby suffer from recall error, are collected in a more systematic manner by government agencies. Fourth, it facilitates a longitudinal research approach over a considerable longer time period than most surveys.

However, administrative registers also have limitations. The two most serious are: first, as it has been collected primarily for administrative purposes, it limits the questions that the researcher can pose, which might tempt the researcher to overlook the most relevant questions and instead only look for questions that can be answered. Second, registers do not include individuals' opinions, attitudes, expectations etc., and thereby pose a challenge for testing different mechanisms. In my opinion, research is not

⁶ The labour market demand for STEM degrees in Sweden has been strong, especially in the field of engineering. Thus, it was and still is common that STEM students are offered jobs even before receiving their degree. All analysis in Study III has also been conducted with only those who have finished their STEM degree, which produces similar results.

⁷ The educational field is collect at age 33 for the youngest birth cohort and at age 42 for the oldest birth cohorts.

⁸ <https://www.scb.se/dokumentation/klassifikationer-och-standarder/svensk-utbildningsnomenklatur-sun/>

only about predicting outcomes, but it is also about understanding why the predicted outcome has come about. One powerful way of understanding why individuals engage in particular behaviors is by simply asking them to provide a reason for their action, and such measures are lacking in administrative registers. However, even given these limitations the above advantages of using administrative registers exceed, in my view, the disadvantages. It is especially fitted for sibling configuration studies, since it enables us to: identify family ties, focus on combinations of sibling configuration factors, and analyze intra-generational outcomes.

The analyses in all three studies in this thesis are based on interaction models. An interaction between two variables (or a two-way interaction) represents a simultaneous association of two independent variables on the dependent variable. In other words, an interaction between two variables means that the association of one of those variables on a third variable is not constant, i.e. the association differs at different values of the second variable. The different characteristics (or variables) that are interacted depend of course on the analysis performed. In order to illustrate this let us take the example of birth order and birth spacing, and their association with income rank. It may then be the case that we expect that first-born siblings on average have higher income rank than second-born siblings, but also that this association may vary by birth spacing. That is, the association of birth order for first-born siblings with shorter birth spacing from their younger sibling might be different from first-born siblings with longer birth spacing from their younger sibling. By interacting birth order and birth spacing with each other, we can evaluate to what extent this is the case.

Analytical strategies when studying sibling configuration

Confounding factors affecting both sibling configuration and sibling adulthood outcomes have been a constant challenge for researchers studying sibling configuration (for an overview see Steelman et al., 2002). The more recent literature in this field suggests that there are possible strategies on how to address this challenge.

One strategy is to use a so-called instrumental variable (IV) that induces a variation in the independent variable but are unrelated to the dependent variable. It is however difficult to find a perfect IV in real life. Two examples of IV used in sibling configuration studies are twin birth (e.g. Rosenzweig and Wolpin, 1980; Black et al., 2005), and miscarriage between two subsequent live births (e.g. Buckles and Munnich, 2012). Neither twin births nor miscarriages are entirely spurious, since they can be induced by parental characteristics or/and behavior. For instance, fertility treatment greatly increases the chance of multiple births (e.g. Fauser et al., 2005), while parents genetics and characteristics, such as high parental age or high alcohol consumption during the pregnancy, increases the risk for miscarriages (Regan and Rai, 2000; Maconochie et al., 2007).

A second strategy is to use sibling fixed effects, which eliminates time-invariant unmeasured factors shared by the siblings. This method has been frequently used for studying birth order effects (e.g. Härkönen 2013; Barklay 2016). However, when applying sibling fixed effects on analyses of birth spacing (Barklay and Kolk, 2017), only sibling groups of three or more children can be studied. Even if these findings are interesting and relevant it is important to note that these families only include a minor share of sibling groups in Sweden. The reason for why two-child families cannot be studied is that fixed effect models requires that there is a variation in the independent variable, and since siblings in two-child families have the same birth spacing interval, this requirement is not fulfilled.

A third possible strategy and the one suggested in this thesis is to use population data and try to restrict the analytical population in a way that eliminates as much heterogeneity as possible, and add control for confounders. The following section therefore describes how the analytical population in the included studies has been restricted in order to adjust for possible confounding factors.

Swedish administrative registers have, as noted above, the advantage of covering all individuals living in Sweden and include information on their biological as well as their adoptive parents (given that they also lived in Sweden), which in turn provides a great ground for creating restricted family ties. The analytical population of Studies I and II has been created to include siblings in two-child families born between 1960 and 1970. The parents are restricted to being biological parents only to the two siblings, i.e. they do not have additional children with another partner.⁹ Further, only intact families have been included in the analytical population, i.e. both siblings have to be living with both parents from age eight to sixteen.¹⁰ The advantage of restricting the analytical population in this way is to ensure that the siblings have had as similar nurturing environment as possible, but also that parental resources (i.e. time, attention, money, etc.) are devoted to only two children, which in turn increases the internal validity of the findings. However, per

⁹ Andersson et al. (2006) show in their study of second- and third-birth risk, that Swedish mothers are more likely to continue childbearing when they have had two same-sex children. Thereby, there is indication that mothers who continue childbearing have a preference for having at least one child of each sex. Thereby, parents in two-child families in Sweden are prone to be more content with the gender composition of their children, or else they would have continued their childbearing. Thus, it can be argued that parental gender preferences for children is not a salient factor for two-child families compared to larger family sizes.

¹⁰ The reason for starting from the age of 8 is due to data restriction. The residential building number (fastighetsnr) is available only starting in 1968; the first sibling would then be 8 years old. It should be acknowledged that living in the same fastighetsnr does not necessarily mean that all family members are living together – that there is a theoretical possibility that the parents (or one of the siblings) could live in different apartments within the same building. However, the probability that two parents who have had children together would be living in the same residential building but in different apartments from the time their eldest child is 8 years old until their younger child is 16 years old should be assumed to be extremely low.

default, when trying to increase the internal validity of the results, the external validity decreases. This in turn means that the result of Study I and II only can be generalized to individuals born between years 1960 and 1970 in Swedish two-child families, who have been living with both of their biological parents until at least 16 years of age.

In addition the analysis also controls for parental characteristics that may work as confounders between sibling configuration and adulthood outcomes, e.g. parental educational attainment, household disposable income, and parental occupation

The analytical population in Study III differs slightly from the above, mainly due to the change of outcome variable from income to educational field. By focusing on educational choices, more specifically STEM fields measured at age 32 or above, I extend the period to include individuals born between 1960 and 1980. Another difference is that only second-born siblings are included in the analyses, and since the focus is on the downstream influence of older sibling to the younger all different two-child families are included. That is, instead of restricting the analytical population to only include biological siblings living in intact families as in Studies I and II, I control for the family type and for how long siblings have lived with their parents. Moreover, since the focus in Study III is to analyze to what extent younger siblings are influenced by their older siblings' educational choices, only sibling groups where the older sibling has acquired his/her degree first are included. In order to take into account for confounding factors, such as parental influences, we include controls for both parents' education level, occupation and whether they have a STEM degree. Individual characteristics, such as living with both parents, GPA score and birth spacing interval.

Overview of the studies

Despite a long tradition, sibling configuration studies have except for a few exceptions paid very little attention to studying the combined effect of birth order, birth spacing sex composition on adulthood outcomes. This thesis therefore contributes to the field by focusing not only on the association between birth spacing and income rank in two-child families, but also on the combined effect of several sibling configuration factors, i.e. birth order, birth spacing, and sex-composition on income rank and educational choices.

Study I focuses on the association between birth spacing and long-run income rank in adulthood among siblings who grew up together in two-child families. By disentangling birth order from birth spacing, the study analyses whether different birth spacing intervals have the same association with income rank for first- and second-born children. The aim of the study is to empirically estimate the income rank gap between siblings with short and long birth spacing, but also to analyze to what extent first- and second born

siblings benefit from different birth spacing intervals. The results show that at the more common spacing intervals (less than 5 years), spacing has a negligible association with second-born children's long-term income rank. However, first-born children suffer when a younger sibling is born when they are very young. Moreover, for both first- and second-born siblings relatively large spacing intervals (over 5 years) are somewhat more associated with lower long-term income-rank, compared to mid-length spacing intervals.

Study II analyses how the addition of another complexity in the sibling configuration, the sex-composition of siblings, is associated with long-run income rank of siblings. The approach is explorative, and the aim of the study is to increase our understanding of the interplay of birth order, birth spacing and sex-composition among siblings in two-child families. By controlling for potential confounders, such as parental education, occupation and parental resources during the siblings childhood, the results indicate clear differences in income rank between the different combinations of sibling configuration factors. The findings suggest that both birth order and birth spacing are important factors for first born boys independent of the younger sibling's sex. First-born girls, however, only have an advantage if they have a younger sister. More surprisingly is that this advantage does not seem to vary by birth spacing.

Study III deviates from the other two studies by focusing on another important stratification outcome, namely educational attainment. Recently, research has started to focus on siblings' configuration and their horizontal tertiary educational choices (e.g. Gabay-Egozi et al., 2016; Barclay et al., 2017). The purpose of this study is to increase our understanding of how younger sibling's choice of STEM degree is influenced by their older brother or sister in Swedish two-child families. The results indicate that younger siblings, net of parental characteristics, are more likely to choose a STEM field if their older sibling already has attended a STEM program. We also find a gender difference among younger siblings choice of STEM fields. The findings indicate that boy's choice of STEM fields seems to be independent of having an older brother or sister who has attended a STEM program. Girls are however more likely to choose a STEM-field if they have a sister who attended a STEM program, than if they have a brother who attended a STEM program. Given that STEM-fields are markedly male dominated, this may indicate the importance of having a same-sex role model for making gender atypical educational choices.

References

- Amin V. (2009) Sibling sex composition and educational outcomes: A review of theory and evidence for the UK. *Labour* 23: 67-96.
- Andersson G, Hank K, Rønsen M, et al. (2006) Gendering family composition: Sex preferences for children and childbearing behavior in the Nordic countries. *Demography* 43: 255-267.
- Ansbacher HL and Rowena R. (1956) The individual psychology of Alfred Adler: a systematic presentation in selections from his writings.
- Bandura A. (1977) Social learning theory Englewood Cliffs, NJ: Prentice-Hall.
- Bandura A. (1986) *Social foundations of thought and action: A social cognitive theory*: Englewood Cliffs, NJ, US: Prentice-Hall, Inc.
- Barclay K. (2014) The Long-term Impact of Birth Order on Health and Educational Attainment. Department of Sociology, Stockholm University.
- Barclay K, Hällsten M and Myrskylä M. (2017) Birth order and college major in Sweden. *Social Forces* 96: 629-660.
- Barclay KJ and Kolk M. (2017) The long-term cognitive and socioeconomic consequences of birth intervals: a within-family sibling comparison using Swedish register data. *Demography* 54: 459-484.
- Bayer AE. (1967) Birth Order and Attainment of the Doctorate a Test of Economic Hypotheses. *American journal of sociology* 72: 540-550.
- Björklund, A., Jäntti, M., & Lindquist, M. J. (2009). Family background and income during the rise of the welfare state: brother correlations in income for Swedish men born 1932–1968. *Journal of Public Economics*, 93(5-6), 671-680.
- Black SE, Devereux PJ and Salvanes KG. (2005) The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics* 120: 669-700.
- Blake J. (1981) Family size and the quality of children. *Demography* 18: 421-442.
- Blake J. (1986) Number of siblings, family background, and the process of educational attainment. *Social Biology* 33: 5-21.
- Blau PM and Duncan OD. (1967) The American occupational structure.
- Buckles KS and Munnich EL. (2012) Birth spacing and sibling outcomes. *Journal of Human Resources* 47: 613-642.
- Bussey K and Bandura A. (1999) Social cognitive theory of gender development and differentiation. *Psychological review* 106: 676.
- Butcher KF and Case A. (1994) The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics* 109: 531-563.

- Böhlmark A and Lindquist MJ. (2006) Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *Journal of Labor Economics* 24: 879-896.
- Caprile, M., Palmén, R., Sanz, P., & Dente, G. (2015). Encouraging STEM studies for the labour market. *Directorate General for Internal Policies, European Union*.
- Carlsson F, Johansson-Stenman O and Martinsson P. (2007) Do you enjoy having more than others? Survey evidence of positional goods. *Economica* 74: 586-598.
- Charles M and Bradley K. (2009) Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American journal of sociology* 114: 924-976.
- Conley D. (2000) Sibship sex composition: Effects on educational attainment. *Social science research* 29: 441-457.
- Conley, D., & Glauber, R. (2008). All in the family?: Family composition, resources, and sibling similarity in socioeconomic status. *Research in Social Stratification and Mobility*, 26(4), 297-306.
- Downey DB. (1995) When bigger is not better: Family size, parental resources, and children's educational performance. *American Sociological Review*: 746-761.
- Erikson R and Jonsson JO. (1996) Explaining class inequality in education: The Swedish test case. *Can education be equalized*: 1-63.
- Evertsson M. (2006) The reproduction of gender: housework and attitudes towards gender equality in the home among Swedish boys and girls. *The British journal of sociology* 57: 415-436.
- Fausser BC, Devroey P and Macklon NS. (2005) Multiple birth resulting from ovarian stimulation for subfertility treatment. *The Lancet* 365: 1807-1816.
- Gabay-Egozi, L., Grieger, L. & Nitsche, N. (2016). Does the Likelihood that a Younger Sibling will Pursue a STEM Major Depend on the Gender of Other Siblings? *Manuscript presented at American Sociological Association 2016*
- Galton F. (1874) *English Men of Science: Their Nature and Nurture*. MacMillan and Company, London
- Haider S and Solon G. (2006) Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* 96: 1308-1320.
- Hauser RM and Kuo H-HD. (1998) Does the gender composition of sibships affect women's educational attainment? *Journal of Human Resources*: 644-657.
- Hauser RM and Sewell WH. (1985) Birth order and educational attainment in full sibships. *American Educational Research Journal* 22: 1-23.
- Hjalmarsson, S., Mood, C. & Shahbazian, R. (2015). *Kodbook för Inkomstvariabler 1968-89 (In English: Codebook for income variables 1968-89)*. The Swedish Institute for Social Research

- Hoem JM. (1993) Public policy as the fuel of fertility: effects of a policy reform on the pace of childbearing in Sweden in the 1980s. *Acta Sociologica* 36: 19-31.
- Härkönen J. (2013) Birth order effects on educational attainment and educational transitions in West Germany. *European sociological review: jct027*.
- Kantarevic J and Mechoulan S. (2006) Birth order, educational attainment, and earnings an investigation using the PSID. *Journal of Human Resources* 41: 755-777.
- Kidwell JS. (1981) Number of siblings, sibling spacing, sex, and birth order: Their effects on perceived parent-adolescent relationships. *Journal of Marriage and the Family*: 315-332.
- Kuo H-HD and Hauser RM. (1997) How does size of sibship matter? Family configuration and family effects on educational attainment. *Social science research* 26: 69-94.
- Maconochie N, Doyle P, Prior S, et al. (2007) Risk factors for first trimester miscarriage—results from a UK-population-based case-control study. *BJOG: An International Journal of Obstetrics & Gynaecology* 114: 170-186.
- Mann A and DiPrete TA. (2013) Trends in gender segregation in the choice of science and engineering majors. *Social science research* 42: 1519-1541.
- McHale SM, Updegraff KA and Whiteman SD. (2012) Sibling relationships and influences in childhood and adolescence. *Journal of Marriage and Family* 74: 913-930.
- Nermo M. (1999) Structured by gender: Patterns of sex segregation in the Swedish labour market: Historical and cross-national comparisons. Institutet för social forskning (SOFI).
- Parcel TL and Menaghan EG. (1994) *Parents' jobs and children's lives*: Transaction Publishers.
- Pettersson-Lidbom P and Skogman Thoursie P. (2009) Does child spacing affect children's outcomes? Evidence from a Swedish reform. Working paper//IFAU-Institute for Labour Market Policy Evaluation.
- Powell B and Steelman LC. (1989) The liability of having brothers: Paying for college and the sex composition of the family. *Sociology of Education*: 134-147.
- Powell B and Steelman LC. (1993) The educational benefits of being spaced out: Sibship density and educational progress. *American Sociological Review*: 367-381.
- Raab, M., Fasang, A. E., Karhula, A., & Erola, J. (2014). Sibling similarity in family formation. *Demography*, 51(6), 2127-2154.

- Regan L and Rai R. (2000) Epidemiology and the medical causes of miscarriage. *Best practice & research Clinical obstetrics & gynaecology* 14: 839-854.
- Rodgers JL. (2001) What causes birth order–intelligence patterns? The admixture hypothesis, revived. *American Psychologist* 56: 505.
- Rodgers JL, Cleveland HH, van den Oord E, et al. (2000) Resolving the debate over birth order, family size, and intelligence. *American Psychologist* 55: 599.
- Rosenzweig MR and Wolpin KI. (1980) Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica: journal of the Econometric Society*: 227-240.
- Shahbazian R. (2017) Is there any difference between having a brother or having a sister? The association between sex-composition and socioeconomic outcomes in Swedish two-and three-child families. *Journal of Family Studies*: 1-27.
- Steelman LC, Powell B, Werum R, et al. (2002) Reconsidering the effects of sibling configuration: Recent advances and challenges. *Annual Review of Sociology* 28: 243-269.
- Sulloway FJ. (1996) *Born to rebel: Birth order, family dynamics, and creative lives*: Pantheon Books.
- Webber DA. (2014) The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors. *Labour Economics* 28: 14-23.
- Zajonc RB. (1976) Family configuration and intelligence: Variations in scholastic aptitude scores parallel trends in family size and the spacing of children. *Science*.
- Zajonc RB and Markus GB. (1975) Birth order and intellectual development. *Psychological review* 82: 74.
- Åslund O and Grönqvist H. (2010) Family size and child outcomes: Is there really no trade-off? *Labour Economics* 17: 130-139.

