Family First? Genetics, within-family differences, and social inequality

Martin Arstad Isungset

UiO : University of Oslo

Dissertation submitted for the degree of Philosophiae Doctor (PhD)

June 2021

Department of Sociology and Human Geography – Faculty of Social Sciences
University of Oslo

© **Martin Arstad Isungset, 2021**

*Series of dissertations submitted to the
Faculty of Social Sciences, University of Oslo
No. 870*

ISSN 1564-3991

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

“Sooner or later,
something that *really* is true will get confirmed,
time after time after time.
And something that in good faith was *thought* to be true,
but isn’t,
when the scientific process repeats it over and over again,
all of a sudden,
you realize,
you know
‘There was something about that that wasn’t quite right.’
So as long as science is humble enough,
and open enough,
and transparent enough to accept the self-correction,
it’s a beautiful process”.

Anthony Fauci,

June 2020,

as depicted in Apollo’s Arrow (2020) by Nicolas Christakis, p. 292

Table of Contents

List of Papers.....	iii
Summary	v
Acknowledgments	ix
1. Introduction.....	1
2. Theoretical Framework.....	7
Counterfactual Causality	7
Genetics and Sociology	9
Sociogenomics, Education, and Income	12
Number of Siblings and Sociology	16
Birth Order and Education	17
Socio-Demographic Trends and Welfare State Norway	19
3. Methodology	23
Estimands: Connecting Statistical Evidence to Theory	23
Twin Methods	30
Genome-Wide Association Studies and Polygenic Scores	32
Population Stratification, Family Trios, and Genetic Nurture	34
Family-Level Fixed Effects and Birth Order	37
Data Sources.....	38
4. Summaries of the Five Papers.....	43
Paper 1	43
Paper 2.....	45
Paper 3.....	47
Paper 4.....	49
Paper 5.....	51
5. Discussion	55
Why Genetics and Nonshared Environments?.....	60
The Future is Now	63
References	67

List of Figures and Tables

Figure 1 - Estimands Framework	24
Figure 2 – Potential Confounding from Genetics	26
Figure 3 - The Classical Twin Model.....	30
Figure 4 - Genome-Wide-Association Studies and Polygenic Scores	33
Figure 5 - Genetic Nurture	36
Figure 6 - Directed Acyclical Graph - Paper 1.....	44
Figure 7 - Directed Acyclical Graph - Paper 2.....	46
Figure 8 - Directed Acyclical Graph - Paper 3.....	48
Figure 9 - Directed Acyclical Graph - Paper 4.....	50
Figure 10 – Directed Acyclical Graph - Paper 5	52
 Table 1 - Research Design of Papers	 28

List of Papers

Paper 1:

Isungset, Martin Arstad, Tina Baier, and Torkild Hovde Lyngstad. 2021. “Genetic Influences on Lifetime Income Increases with Gender Equality”. Submitted to *American Sociological Review*. Available at *SocArXiv*, doi: 10.31235/osf.io/e5wjk

Paper 2:

Isungset, Martin Arstad, Dalton Conley, Henrik D. Zachrisson, Eivind Ystrøm, Alexandra Havdahl, Pål R. Njølstad, Torkild H. Lyngstad. 2021. “Social and Genetic Effects on Educational Performance in Early Adolescence”. Ready for submission, available as *NBER Working Paper* (28498):1-27

Paper 3:

Isungset, Martin Arstad; Mats Lillehagen; and Elisabeth Ugreninov. 2020. “One Order Fits All? Birth Order and Education in Immigrant Families.” Published in *European Sociological Review* 36(1): 16-31

Paper 4:

Lillehagen, Mats, and Martin Arstad Isungset. 2020. “New Partner, New Order? Multipartnered Fertility and Birth Order Effects on Educational Achievement.” Published in *Demography* 57(5):1625–46

Paper 5:

Isungset, Martin Arstad, Jeremy Freese, and Torkild Hovde Lyngstad. 2021. “Birth Order Differences in Education Are Environmental in Origin”. Submitted to *Proceedings of the National Academy of Sciences*. Available at *bioRxiv*, doi: <https://doi.org/10.1101/2021.06.0.446519>

Summary

Social inequality is one of sociology's core research areas. Much sociological research implicitly assumes that the causes of variation between individuals in socioeconomic outcomes are environmental in origin, often emphasizing the individual's family environment as a main source of differences. This thesis explores this tacit assumption by answering two overarching research questions: The first question asks what can we learn from including genetics in studies on social inequality. Genetics is often put forward as a potential confounder for sociological explanations emphasizing social background effects as being environmental in origin. From another vantage point, the second research question asks what can we learn from including nonshared environmental influences, like birth order, in studies of social inequality. Should we expect siblings to have similar socioeconomic outcomes or are sibling dissimilarities overlooked in social inequality research.

The thesis consists of five papers and one overarching introduction. The first two papers use genetically sensitive designs and considers whether the relationships between social background and socio-economic outcomes are confounded by genetics, using education and income as case studies. In paper 1, we use administrative register data and look at lifetime income over birth cohorts 1915-1991 in Norway using twin methods. We find that, looking past differences due to gender, shared environmental influences which encompass social background variables such as parental educational attainment and - income, do not influence lifetime income to a large degree. Instead, genetic and nonshared environmental influences are high. Which influences that are the most influential have changed over birth cohorts. In particular, they have changed in relation to gender equalization, where women faced considerable structural constraints to earn income earlier in the 20th century. For income, we learn that shared environmental aspects of social background that are not related to gender have less of an influence, and that genetic- and environmental influences twins do not share are more important.

In paper two, we use administrative data in combination with molecular genetic data from a large sample of genotyped trios of mothers, fathers, and offspring. We investigate the intergenerational transmission of education through environmental and genetic pathways. Using polygenic scores (PGSs) for both parents and offspring, and parental educational attainment, we consider whether the social background effect of parental educational attainment is confounded by genetics. Our outcome is educational achievement in 5th, 8th, and

9th grade. Parental educational attainment and own polygenic score both influence educational achievement, and are similar in magnitude. They are largely independent of each other, and social background effects are confounded only to a small degree. However, we also investigated whether parental genetics influence offsprings' educational outcomes above and beyond offsprings' own PGS. There are effects of untransmitted parental genotypes, a phenomenon known as “genetic nurture”. This shows that genetics are involved in social background effects as parents' genetics influence parents to provide advantageous environments for educational achievement for their offspring

In papers 3-5, and indirectly in paper 1, we consider sibling differences. Papers 3 to 5 are on birth order differences. In paper 3 we investigate birth order differences for educational achievement in adolescence and whether these differ across immigrant- and native families. We use comprehensive administrative register data, allowing us to break the sample into subgroups according to parental country of origin. Similar to what other research finds, we find that earlier-born have higher educational achievement than later-born in native families. By and large, we find the same for immigrant families with various parental countries of origin, but for some origins, there are no birth order differences. This is surprising in light of the main theories on birth order differences which typically are universal, and we theorize that it may due to differences in parenting strategies, child care attendance, and peer influence. Notwithstanding, from this paper, we learn that sibling differences in education due to birth order are pervasive, and exist in families with different cultural backgrounds from native populations.

In paper 4, the outcome is also educational achievement in adolescence, and we again make use of administrative register data. We analyze birth order differences in families where the parents have offspring with several partners. We label two different birth orders; full biological birth order which is your birth order considering full siblings only, and overall birth order which includes all full- and half-siblings. Birth order differences exist in these cases too, which is not obvious as separation and divorce typically have a negative effect on offsprings' educational achievement. We find that the influence of the two different birth orders is connected to the sex of the multipartnered parent. In the mother-based sample, there are overall birth order differences. In the father-based sample, only full biological birth order differences are found. We believe this to be a result of the residential arrangements following separation, or mothers being more involved in childrearing in general across sibships. From

this paper, we once more show the pervasiveness birth order differences, demonstrating the importance of this within-family inequality-generating mechanism.

In paper 5, we consider the potential genetic confounding of birth order differences. We use administrative data, survey data, and molecular genetic data from trios of parents and offspring. First, we look at PGSs and whether they vary with birth order, which they do not. Second, we consider in utero environments, proxied by birth weight and birth length. Later-born are heavier than earlier-born, and birth weight is often found to be positively associated with socioeconomic attainments later in life, meaning that later-born siblings are born with what is statistically an advantage. Third, we consider educational outcomes. For educational achievement in adolescence, earlier-born children perform better than later-born children, taking their genetics and in utero-variables into account. Looking also at educational attainment and birth order in the parental generation, we find that PGSs does not influence educational attainment. From this paper, we once more learn about the pervasiveness of nonshared environments such as birth order, and the key takeaway is that birth order differences arise early in childhood due to environmental influences.

To sum up, this dissertation shows that including genetics and nonshared environments like birth order offer a broader perspective to what causes social inequalities. Sociology has by and large ignored genetics and within-family differences, both of which can be larger than between-family environmental influences on outcomes of interest to social inequality scholars. Going beyond the tacit assumption of homogenous family environments as the main mechanism behind social inequalities, sociology may find that many socio-economic outcomes are more structured than currently acknowledged. The first two papers show that genetics confound social background perspective to a substantial degree for income, but to a lesser degree for education. This fits well with findings from earlier genetically sensitive research. The three last papers show that we should not expect siblings to be similar in educational outcomes, they are dissimilar according to their birth order.

Acknowledgments

Torkild Hovde Lyngstad, you are an amazing person, scientist, and sociologist, in no particular order. You are always available and have solutions to every problem, even if you're the busiest and most productive person I know, and you have a family life. You have contributed greatly to all papers and supported me financially to go to workshops and research stays. You have included me in networks and provided data. You've coded and debugged. We've met in the holidays. You even let me stay with you and get to know your family at Berkeley. Thank you for being a role model and a super nice guy I'm proud to affiliate myself with.

Mats Lillehagen, my birth-order-brother. I'm grateful for our time working together, which laid the ground and built my confidence for the rest of this thesis. Your openness to all sorts of ideas, arguments, and knowledge in a bewildering myriad of topics is something I cherish. I'm thankful that our friendship is not solely based on discussing scientific topics, an important feature to my well-being at work.

Jeremy Freese, our conspiring on DNA was hampered by a piece of RNA sweeping the world. I'm very grateful that you continued to work with me via your local providers despite the enormous hurdles 2020-21 posed. I hope your creativity, curiosity, positivity, and grit are as contagious as COVID-19, and that I'll "suffer" long-Jeremy symptoms in the years to come.

Elisabeth Ugreninov, thank you for your supervising my master's thesis and for pushing and helping me with my two first publications. I learned a lot from your ideas and research skills for writing those papers and I would've never embarked on it without your support.

I would like to thank all my other co-authors in this thesis for improving the papers beyond what I could've possibly done on my own. Thank you Tina Baier, Dalton Conley, Alexandra Havdahl, Henrik D. Zachrisson, Eivind Ystrøm, and Pål R. Njølstad.

All previous and current members of Social Inequalities and Population Dynamics (SIPD) deserve a big thanks for all the support and for providing a very safe environment to grow in as a junior researcher. I may be somewhat biased and I know you generally don't like bias, but to me, you are truly the best sociological research group period. A special thanks to our "Mitochondrial-Eve", Gunn Birkelund. You're an incredible researcher and resource that have benefitted so many (young) scholars, and I'm no exception. A big shout out to the OPENFLUX-team: Torkild, Tina, Agnes, and Ruth, for creating a social science genetics

environment. Thanks to the Oslo Meeting for Social Science Genetics (OMGSS) milieu. More to come soon for all groups I hope.

Thank you to Uzair Ahmed and Inga Sæther for making sharing an office as good as it can get. I look forward to the day we start every conversation with “Doctor Sæther and Doctor Ahmed”.

I want to thank Fulbright, Wedel Jarlsbergs fund, OPENFLUX, and the Department of Sociology and Human Geography for funding me to go to Stanford.

I’m grateful for the opportunity to work at the Department of Sociology and Human Geography, where there are so many skillful and interesting colleagues. A special thanks to Inger-Lise. Effective admin is so important, and you’re the best I’ve experienced.

I’m glad to have people I can call friends inside and outside of academia that gives me valuable relationships, perspectives, and laughs. None mentioned, none forgotten. To my family, Anne Lise, Odd, and Ane; you provided me with the best possible components. I thank you for all your support throughout my life.

Hanne, I can now proudly call you my wife, wife. Thank you for being there for me and all your love and support. I’m especially humbled and grateful that you were in California with me when the world came tumbling down. I don’t know how I would have fared in that situation without you. But I know that I wouldn’t have submitted this thesis by now. Thank you for sharing your life with me.

Martin Arstad Isungset,
Oslo, June 2021

1. Introduction

Social inequality is a central pillar of sociology. From the classical sociological writings by Marx and Weber, through the pioneering quantitative work of Duncan (and Blau), to 1980-2000s class schemes of John Goldthorpe, Erik Olin-Wright, and Pierre Bourdieu, and the latest revitalization with Piketty, it remains pivotal in the discipline and our societies. Throughout the period, research filling countless library shelves have shown that social background influences individuals in their attainments of socio-economic goods. An (implicit) assumption and practice for many scholars working with social inequality are that the social family environment, as measured by parental characteristics such as education, income, and class, influences the socio-economic outcomes of the offspring. Another (implicit) assumption is that family environments influence siblings to be similar, not different. Throughout this thesis, I investigate whether the “default hypothesis” of homogenous influence of family environments is warranted for social inequality issues.

Providing a backdrop, a mass collaboration showed that scientists have a hard time predicting life outcomes even with very rich family data (Salganik et al. 2020). This led to questions being raised whether we are missing essential information on families that need further exploring (Garip 2020). In this thesis, I seek to broaden the scope of how inequality comes about by considering two factors related to families that are often overlooked in sociology. The first factor is genetics. Genetics can underlie differences between families, confounding the default hypothesis of homogenous effects of family environments causing social inequality. Within families, genetics can cause differences too, as non-twin full-siblings share 50 % of their genes.

The second factor is nonshared environments. As shown by sibling correlations, siblings are surprisingly different (Grätz and Kolk 2019; Wiborg and Hansen 2018). As mentioned, sibling differences can be due to siblings sharing only 50 % of their genetic makeup. But sibling differences can also be caused by environmental influences siblings do not share, such as peers, random events or luck, and their place in the family, i.e. their birth order. Together, this puzzle of genetics and sibling differences lead me to ask two overall research questions in this thesis: 1) What can we learn from including genetics in studies of social inequality? 2) What can we learn from including nonshared environmental influences like birth order in studies of social inequality? In the rest of the introductory chapter, I briefly

describe the family environmental paradigm in sociology and neighboring scientific fields that challenge it.

Historically, the explanatory precedence of family environments in sociology has long roots ranging back to the canonical Enlightenment-era-scholars and was perhaps most vehemently found in Durkheim. The family environmental roots of 20th Century sociological quantitative work can be traced to Talcott Parsons. Durkheim, Weber, and Parsons had a dualistic notion of reality (Aakvaag 2010), where the social sciences are different from the other sciences, laying the ground for exclusively social understandings of modern society. With Parsons, structural functionalist theory became paradigmatic in sociology in the 1950s (Parsons 1953). Central to the structural functionalist perspective was the notion of societal order and its maintenance. The AGIL model solved the issue for Parsons (1968), where social order is maintained by external pressure, but importantly also through internalization through socialization. Written at the high-point for of the housewife-family norm in the 1950s Western societies (Klepp 2018), the main arena for internalization of social norms was thought to be the family. This set the stage for sociological thinking about individual socioeconomic outcomes over the coming decades.

Even though influential scholars such as Bourdieu, Foucault, and Giddens pitted their theories of society against the structural functionalist perspective, family socialization remained a central dogma, albeit often implied rather than explicitly part of the theory like under Parsons (Guhin, Calarco, and Miller-Idriss 2021). The social effect of family background remained a “model monopoly” (Bråten 2000) in the period following the fall of structural functionalism as a unifying perspective. In a model monopoly, alternative (research) questions are not asked, and other perspectives are defined as irrelevant. According to Bråten, it’s not necessarily an intentional act, rather the agent may be spellbound by the leading perspective. Other perspectives may not be actively discouraged, but model monopolistic dominant actors engage only in shallow dialogue, as they disregard other perspectives a priori. The horizon of meaning in scientific conversations is hampered and could have been widened instigating better understanding for all involved parties if only the engaged actors realize what they had engaged in. The social family background paradigm in sociology remained, to put it with a classic, the dominant discourse.

The causal precedence of social background was, and still is, widespread in the social sciences. So dominant was social constructivism and the social background paradigm in the social sciences that Steven Pinker subtitled his bestseller where he problematizes this

perspective *The Modern Denial of Human Nature* at the beginning of the 2000s (Pinker 2002). According to Pinker, the social sciences dishevelment from the natural sciences has completely dislodged the social sciences from fundamental facts of the human species. Most social scientists accept modern evolutionary theory. They accept that we are the results of millions of years of evolution following Darwinian principles. Yet, for decades, the social background paradigm was almost completely unaffected by this. It is less dominant today but remains in many branches of sociology. One influential introductory book to sociology used at many undergrad courses across the globe is penned by one of sociology's most prominent still-living scholars, Baron Anthony Giddens (Giddens and Sutton 2017). It contains only one reference to genetics, DNA, biology, or similar terms, except gene therapy and genetically modified foods. The reference is under "intelligence-controversy". It references Charles Murray which in the 1990s posited controversial viewpoints on the link between race, intelligence, and genetics. Thus, I'm not drawing a straw-man here as the 1 200-page strong introductory piece find no place for genetics with exception to the rebuttal of Murray, who is a highly controversial figure in the biological sciences too, by most regarded as an ideologically (mis)guided scholar. Murray as the pick of the litter of genetically oriented social scientists is picking far out in the tails of the normal distribution.

Meanwhile, in a sub-branch of psychology, behavioral genetics, there has been no withdrawing from human nature over the same period, rather the opposite. The core aim of behavioral genetics is to study how genetic differences among people contribute to differences in their behaviors (K Paige Harden 2021; Knopik et al. 2017). Behavioral genetics applies statistical techniques to investigate influence from genetics and environments on any given trait in humans. In a direct line from Darwin and Mendel, to Fisher, Galton, and others, behavioral genetics has a long (and historically problematic (Haave 2000; Mukherjee 2015; Nielsen, Monsen, and Tennøe 2000)) tradition of using human pedigrees to assess genetic and environmental influences on specific outcomes. Sociology is a social science, so it's not imminently clear that sociologists should give behavioral genetics any attention. Perhaps social family background should remain the monopolistic model without further considerations? I would disagree.

The first problem of ignoring findings from behavioral genetics that guides the first research question is confounding. I use twin methods as an introductory conceptual wrapping for this thesis. Typically, the main focal point in behavioral genetics has been the influence of additive genetics, *the heritability coefficient*. The heritability coefficient is denoted A in twin

models. A captures variance that is due to monozygotic (MZ - identical) twins sharing 100 % of their genes, as opposed to 50 % like dizygotic (DZ - fraternal) twins do. In addition to the heritability coefficient, twin models output two environmental components. C captures shared environmental influences that cause twins to be similar. C capture environments that sociologists typically invoke for their explanations, like parental educational attainment, income levels, and family environments, the explanatory forces often invoked in the social background paradigm. Lastly, E captures nonshared environmental influences, environmental influences that cause twins to be different from each other that they do not share, like friends, teachers, luck, and random events. The classical twin model is often denoted as the ACE-model based on the components I have just described. Most commonly, one uses standardized ACE-models which restrict the variance, to sum up to 100 %, making it possible to attribute genetic and environmental influences in percentages, comparing the influence of each component.

The heritability coefficient captured in A then, as opposed to what is captured by shared environments in C, would indicate how much confounding the social family background paradigm is at risk of, as most sociological studies do not distinguish between these two influences. A meta-analysis of 50 years of twin studies shows that across all traits the heritability is 49 %, and “[t]he data are inconsistent with substantial influences from shared environment [...]” (Polderman et al. 2015:702). However, quite a few of the traits considered in the meta-analysis are far from the traits sociologist typically consider, and closer to medical and biological sciences where one would expect some genetic basis for them.

Socio-economic outcomes are also substantially heritable. Nielsen (2018) investigate status-related outcomes with twin models. College plans, occupational wages, educational attainment, household income, verbal intelligence, GPA, and subjective social status all have a high heritability, much higher than influence from shared environments. This finding, that genetic influences seem to dominate shared environmental influences, is so common that Eric Turkheimer famously formulated three laws: 1) All human behavioral traits are heritable; 2) The effect of being raised in the same family is smaller than the effect of genes; 3) A substantial portion of the variation in complex human behavioral traits is not accounted for by the effect of genes or families (Turkheimer 2000). Sociological inquiries of individual-level outcomes then, seem to be in big trouble by confounding. However, Turkheimers laws are formulated as laws to make an important point: genetics is a systematic source of variability,

but for methodological reasons rather than substantive ones. Nonetheless, social family background does not seem to influence siblings in the same family to be similar mainly due to shared environmental reasons, but due to genetic similarity, which poses a puzzle to sociology on how to reconcile this with sociological theories and methods.

The second problem, guiding research question two, is motivated by Turkheimer's third law, that which states that nonshared environmental influences make up a substantial portion of human behavior. This problem is not a challenge only to sociology from behavioral genetics, as it concerns all behavioral sciences to a larger or smaller degree. Our theories and statistical models are not able to predict or explain very well how people will end up (Salganik et al. 2020). For instance, models like sibling correlations are often used to capture the total family effect (Mazumder 2011), all things that make siblings similar to each other in a given outcome. As already discussed, socio-economic outcomes are heritable, and a substantial part of the total family effects is a genetic effect (Freese and Jao 2017; Vladasel et al. 2020), potentially confounding the estimate.

Moreover, sibling correlations (between full siblings) account for shared genetics, but as siblings share on average 50 % of their genes, it does not account for nonshared genetic influences. Social inequality could thus be even more structured than sociologists currently believe. Furthering the depth of the problem, sibling correlations for socio-economic outcomes like education and income are 0,01-0,40 in Norway (Grätz and Kolk 2019; Wiborg and Hansen 2018). For some outcomes like income, different-sex siblings are only marginally more similar than two random people drawn from the population (sibling correlation = 0,01-0,15). A large part of siblings' outcomes is not explained by factors shared by the siblings (Grätz 2018; Plomin and Daniels 1987). Moreover, sibling correlations are descriptive and could arise through many processes (Lundberg 2020). Even with the latest fad of predictive methods like machine learning and very rich historical family data, (social) scientists struggle to predict individual-level outcomes (Salganik et al. 2020). The depressing results from this mass-collaboration of prediction across scientific communities left some to ask whether we are missing essential things on some families and whether there are problems with the data as much as the methods (Garip 2020). I would add to that account that the problem may well reside in our theories and expectations. In particular, I wonder whether the theoretical notion that homogenous effects of (exogenous) social family background are our best bet for predicting sociological outcomes.

As sibling correlations show that siblings can be expected to be quite different, perhaps we should consider that difference more closely? However, sibling dissimilarities like nonshared environmental influences have proven elusive to pin down (Turkheimer 2016). One important exception is differences according to birth order, which is one of the focal points in this thesis. Siblings in the same family are different according to birth order, with earlier-born siblings having higher intelligence (Barclay 2014; Kristensen and Bjerkedal 2007), higher educational attainment/achievement (Barclay 2015; Black, Devereux, and Salvanes 2005; Kristensen and Bjerkedal 2010), and higher income (J. R. Behrman and Taubman 1986) than later-born. Differences according to birth order are large enough to warrant a larger scholarly interest than currently given, and can be of similar magnitude to other social background variables (Härkönen 2014). Black et al (2005) find that the difference between first- and fifth-born in educational attainment in Norway is of similar magnitude as the difference between blacks and whites in the United States. Similarly, Isungset (2015) finds that the difference between first- and fourth-born in lower secondary school grades is similar to the gender gap in the same outcome in Norway. As birth order differences are of considerable magnitude and show that siblings in the same family have different socio-economic attainments, this underlines the problem of assuming that family environments influence siblings in the same way.

The overarching research questions guiding this thesis are: 1) What can we learn from including genetics in studies of social inequality? 2) What can we learn from including nonshared environmental influences like birth order in studies of social inequality? I answer the first question by using two types of genetically informed data, twin data, and molecular genetic data, and administrative registry data. The second question I answer using administrative registry data, and for one paper a combination of register-, survey- and molecular genetic data. In the following chapter, I will describe the theoretical backdrop I use to derive hypotheses to answer both the overall questions in the thesis and the more specific research questions in the papers. In chapter three, I introduce the methods and statistical techniques I use in the papers. In chapter four, I briefly describe the papers' place in answering the overall research questions. Finally, in chapter five, I discuss what we can learn as sociologists by including genetics and nonshared factors like birth order.

2. Theoretical Framework

In this chapter, I will discuss the philosophy of science linked to counterfactual causality that guides the papers and research questions in this thesis. I then pursue this line into broader aspects of how sociology may make sense of genetics and nonshared environments. As each paper contribution has a literature review section relevant for each one, I do not delve into the more substantive theoretical mechanisms and evaluations here.

Counterfactual Causality

The introductory chapter to *The Oxford Handbook of Causation* (Beebe, Hitchcock, and Menzies 2010; Paul 2010) does not contain a definition of causality. Even attempting to define causality philosophically would give one notion (causal) precedence over another. I use Judea Pearl's influential work to rank-order notions of causality as a guide to how causality can be understood. I briefly introduce the two first notions, before focusing on counterfactual notions, which according to Pearl are at the highest rung of this three-rung ladder (Pearl and Mackenzie 2018). At the lowest rung of the ladder is regularity/association theories.

Association theories are law-like causality-theories, which are both retrospective and prospective. A causes B iff 1) A is occurring before B in space and time; 2) B follows A in time; 3) all A events are regularly followed by events B (Psillos 2010). We seldom, iff ever, adhere to regularity theories in the social sciences, as the restrictions are in practice never met for social phenomena.

Rung two are interventionist/agency theories of causality (Woodward 2010), or as Pearl calls them Doing (Pearl and Mackenzie 2018). Doing something to the world and observing whether the outcome changes is the core of this notion of causality. If A causes B, then if A were to be manipulated in the right way, there would be an associated change in B (Woodward 2010). Under regularity theories, correlations between variables could have been observed for all eternity, but are not causal as they have not been manipulated to show that A is causally instigating B. Shoe size can be used as an example, there is a positive correlation between shoe size and crime. However, few would contend that larger shoe sizes cause individuals to commit criminal acts. Rather, many would point to men having larger shoe sizes than females, and men having a higher likelihood of committing and being prosecuted for criminal acts. In practice, interventionist conditions are very hard to do in social sciences.

We cannot swap people's genomes (yet), for instance, to see whether such an intervention changes anything in our chosen outcome.

Pearl's general solution to science is to rely on counterfactual notions of causality, the highest rung on the ladder, which he coins *Imagining*. This word is not often used in scientific writing and may seem like a strange word to use when describing causality. Paul (2010) defines counterfactual causality: A causes B because the counterfactual “if not A, then not B”. Note that this does not mean that B is always caused by A. B could be caused by all sorts of letters. Counterfactual causality is established by imagining what would have been the case for our observed outcome, had another event occurred than the one we believe caused it. It is not intuitive that this is indeed a “better” notion of causality. Why should we accept mere imagination, we (believe we) can imagine anything! Yet, this is the exact reasoning Pearl puts into counterfactuals as having the highest rung of the causal ladder. Using Yuval Harari's example about The Lion Man of Stadel Cave, Pearl claims that this 40 000-year-old sculpture of a creature with a man's body with a lion's head is a precursor of every philosophical theory, scientific discovery, and technological innovation (2018:34–35). Why? Because it depicts an imagined creature allowing us to question what is. Every single philosophical, scientific and technological advance had to be imagined to be put into this world, Pearl reasons, and the rise of science, innovation, and technology is caused by homo sapiens asking the question “Why?”.

Sociologists too ask why-questions, we purport to causal explanations of social phenomena. When sociologists ask “Why do men have higher incomes than women?”, they seek to establish causal explanations to that questions. If not, the question would be descriptive: “Do men earn more money than women?”. Descriptive questions are also, of course, often asked by sociologists. But in social inequality studies, sociologists more often than not seek causal answers, howbeit they may not be explicit about it. For instance, varying parental educational attainment and seeing whether it cause differences in outcomes for children can be regarded as a counterfactual assessment. Counter-to-fact, one poses the question of what would have happened if an individual's parent was more educated as opposed to less educated, and contrast the two using statistical methods. A parent cannot have both high and low educational attainment at the same time. Thus far, sociological accounts of social background differences causing intergenerational effects match counterfactual reasoning. If one is interested in the total effect of having higher educational attainment on the next generation's educational outcome, the model I just described with one X (parental

educational attainment and its effect on Y (children's educational outcomes) is what you are looking for. However, sociologists typically fill in the gap between what could cause a relationship between parental educational attainment and children's educational outcomes with sociological theory including family environments. In principle, another variable is added to the theoretical model, but not tested, only theoretically inferred. To be clear, if only sociologists were precise about what they are doing (Lundberg, Johnson, and Stewart 2021), adding another parameter to the causal model, this would be in perfect conjuncture with what Pearl would like scientists to do. Clear assumptions make for clear(er) testing of these assumptions. Adding another variable allows for testing that variable empirically, but sociologists by and large do not include family environments, but infer that there is a social transmission of educational background going on, seldom paying confounding from genetics any mind.

Genetics and Sociology

Released in 1967, Peter Blau and Otis Dudley Duncans *The American Occupational Structure* (1967), is a landmark in research on social inequality. Blau and Duncan sought to investigate how intergenerational transmission of occupational status attainment happened between fathers and sons. Did the path from Origin (e.g. parental occupational status) go through intermediate traits such as offspring Education, and then on to the offspring Destination (i.e. son's occupational status)? Or did it go directly from Origin (parental occupational status) on to the Destination (son's occupational status)? In the book, the authors use structural equation models (SEMs). The SEMs Blau and Duncan used were intended to display causal relationships (Pearl and Mackenzie 2018:285), a theoretical causal model for investigating intergenerational social inequality. Model parameters from SEMs, however, are only informative if the underlying theoretical model is correct. Having read Eckland (1967), Duncan¹ (Duncan 1968; Nielsen 2006, 2018) admitted that the findings in their book did not sufficiently investigate merit and the potential genetic component in intelligence that could confound the paths in their model. In this later paper, he re-did some of the analysis done in the book and found that intelligence “serves as a kind of springboard, launching many men into achievements removing them considerable distances from the social class of their birth”.

¹ Duncan published in a journal which at the time was named “Eugenics Quarterly”. I leave it to the reader what to make of that.

The sociological theories on social inequality that dominated following Blau & Duncans 1967-classic for the most part chose to ignore the possible confounding from abilities running genetically in families and focused instead on purported homogenous influences from social background through family environments.

Granted, genetic and social data were harder to come by than just social data as they could be more easily collected. Be that as it may, many sociologists chose to form their theories to focus less on abilities and genetic influences, with notable exceptions like the proponents of the so-called Wisconsin Model who included “mental ability” in their models (Sewell, Haller, and Portes 1969), and scholars in the milieu associated with Christopher Jencks (Jencks et al. 1972). The lack of genetic anchoring has consequences for causal interpretations still prominent in sociology today. Whether one admits it or not, just leaving genetics out of the model does not alleviate the concern that your parameter estimates may be confounded. Again, if one is interested in the total effect of parental origins on child destination, it’s totally fine to leave genetics out. However, as many of the sociological theories attribute this to social background, it remains a causal identification problem without removing potential genetic confounding. In other words, we would need to counterfactually imagine how genetics could play a role in social inequality and then attempt to empirically assess that counterfactual notion.

How, then, should sociologists understand genetics and social inequality? Freese (2008) offers a counterfactual interpretation to quantitative sociologists. The question “Who completes higher education?”, is often translated as “What is the effect of x on y ?”, where x could be parental educational attainment. Freese envisions the counterfactual assessment of swapping individuals' genetics. One could swap the total genome of a person with a different one, or one could swap single genes or one could swap parts of genomes. We can then ask questions like “Who completes higher education?”, and counterfactually swap genomes to see whether some genomes causally influence some people to complete more than others. According to Freese, the results from behavioral genetics that almost all human traits are substantially heritable (Polderman et al. 2015) should lead sociologists to expect the *ubiquitous partial heritability thesis* (UPHT): genetic differences are partial causes of the overwhelming majority of individual-level outcomes of interest to psychology and social science. Accepting the UPHT is the first step.

Conceptually, twin studies are based on counterfactual reasoning. Closely related to path models and often estimated using structural-equation models, they are very clear on the

assumptions made (see next chapter). Heritability estimates, and the shared and nonshared environmental components that are estimated with them in twin methods, are population estimates of what causes differences between people. Being population estimates, they can vary with time, historical contexts, institutional arrangements (Knopik et al. 2017), thus “the biologist’s trash has turned out to be the psychologist’s and sociologist’s treasure” (K Paige Harden 2021:1217). This is one of the reasons why I use twin methods and the ACE components as a conceptual wrapping in this thesis. Heritability, however, is often misunderstood. The heritability of a given trait is sometimes interpreted as genetics determining the trait. That’s wrong. The heritability of having a nose in the human population is 0. It’s 0 because almost all humans have a nose, there is not enough variation in having a nose or not in the population. Heritability and the corresponding environmental components do not show how much genetics determine a given trait, but how much of the variation in a trait genetic influences explain in the current population at the given time of measurement. However, should we expand our population to include humans and other species that do not have a nose, the heritability of having a nose would be something (Sesardic 2005).

Some have proposed that we can use such estimates as roughly analogous to Gini-coefficients for causes of social inequality, they show us how the outcome of the genetic- and social lottery differs across time and place (K Paige Harden 2021). As an illustration, total family effects could be fully confounded by genetics, or fully determined by family environments without any confounding, and the social science scholar applying them could not tell the difference as the different components would be masked (Branigan, McCallum, and Freese 2013). Per contra, heritability estimates from twin studies would tell you that the variation is linked to genetic differences or shared environmental differences. As the genetic consequences of the evolution of the human species are happening at a very slow rate (Kong et al. 2017), genetic associations may help us interpret social change, as genes are close to static, whilst environments may change drastically over very close time frames. Circling back to Freese (2008), we do not get closer to the actual mechanism by accepting the UPHT, but by using genetically sensitive design in conjuncture to varying with the meso- and macro-level institutions sociologist care so much about, we learn more on how and when genetics, shared environment, and nonshared environments matter, and if they change when structures, norms, and possibilities change. Some steps that can be made to getting closer to mechanisms free from genetic confounding are using institutional variation to investigate such *gene-environment interplay*, gene-environment correlation and -interaction, finding natural

experiments as instrumental variables (IV), or using individuals as their own controls. The relationship between genetics and different environments, Freese argues, is more about the relationship between individuals and social structures, not about genetics per se. In other words, it's akin to the agency-structure problem sociology so often finds itself in and can offer a promising avenue for unlocking parts of this problem (Mills and Tropf 2020). Of course, we would still need ample sociological theory and imagination to describe this relationship, which we by and large lack today.

In the social sciences there is a small, but rapidly growing, literature considering how genetic influences vary with institutions, historical contexts, gender, and other environments, sometimes placed under the umbrella terms sociogenomics or social science genetics (Conley and Fletcher 2017; Freese 2018; Harden and Koellinger 2020; Mills and Tropf 2020). Pedigree-based studies like twin studies, adoption studies, sibling studies, and parent-offspring correlations were for long the main mode of investigating potential genetic influences. Since the sequencing of the human genome 20 years ago (Lander et al. 2001), molecular genetic approaches have come into use. The decreasing cost of genome sequencing has led to massive improvements in data availability and many important methodological advances. I focus this brief walk-through on the studies which use the two most widely applied genetically sensitive designs, classical pedigree-oriented studies and studies applying polygenic scores (PGS), and focus on studies with gene-environment interplay. As the papers in this thesis consider education and income as outcomes, I focus on these and other SES-related outcomes in the following sections.

I describe PGS in detail in chapter 3. For this chapter, we can think of PGSs as capturing variation in the DNA of an individual, for the outcome of interest, and their most important feature is that they work (Freese 2018), i.e. they have some predictive power that stems from genetic differences between individuals.

Sociogenomics, Education, and Income

Education is one of the most studied sociogenomic outcomes due to its availability in datasets with genetic information. Most surveys ask for the educational attainment of the respondent regardless of study aim, enabling researchers to use education as an outcome even for datasets originally collected for other purposes. Education of course follows the First law of behavioral genetics (Turkheimer 2000), it's substantially heritable. However, educational

attainment is one of very few human traits that show a considerable magnitude of shared environmental influences too (Branigan et al. 2013; Nielsen and Roos 2015), which is puzzling (Freese and Jao 2017). An influential meta-analysis found that the pooled heritability of educational attainment is 40 %, and pooled shared environmental influence is 36 % (Branigan et al. 2013). The results were similar in a more recent meta-analysis (Silventoinen et al. 2020). For education, sociologists' fears of confounding are therefore less of a concern than other outcomes. Nonetheless, heritability is also of considerable magnitude for educational attainment, and proposed homogenous effects of total family background could still be severely confounded by genetics.

Gene-environment interplay, like counterfactual varying of macro-level institutions to see if genetic and environmental components change, have shown interesting results. Findings seem to support the idea that heritability is higher in societies that provide more equal opportunities (K Paige Harden 2021). Heath et al. (1985) use data on educational attainment for Norwegian twins over several birth cohorts (1915-1960), and find that as liberal and social policies come into place (their interpretation), heritability rises and shared environmental influences decreases. The authors interpret high heritability as signifying higher merit, however, the mechanisms are not clear and others contend that we should interpret rising heritability with caution until we know how heritability manifests itself (Freese 2008). Similar results to Heath et al. (1985) have been found in Sweden too (Lichtenstein, Pedersen, and McClearn 1992).

Engzell & Troup (2019) show that in societies that have high intergenerational educational mobility as measured by parent-offspring correlation, shared environmental influence is lower and heritability is higher. This indicates that higher heritability can be used as a Gini-index of what causes social inequality (K Paige Harden 2021) and that we can expect higher heritability in societies that provide more equal opportunities for all. Findings from twin studies are corroborated with studies using molecular genetic data. A study based on data from Estonia found that the educational attainment PGS has a higher influence on educational attainment for birth cohorts that came of age after the Soviet Union collapsed than those who came of age before it collapsed (Rimfeld et al. 2018). When Estonia became a more open and democratic society following the collapse, genetics matters more than when it was under a communist centralized regime.

About the intergenerational relationship between genetics, and social background for education, Conley et al. (2015) and Liu (2018) use sociogenomic data from the USA. Using a

PGS for educational attainment, both studies find that the confounding of the social background effect of parental educational attainment is small, $1/6^{\text{th}}$ and $1/5^{\text{th}}$ respectively. Genetics and social background both play a role but are largely independent of each other. Besides, both studies use parental genetics to see whether there are genetic effects over and above children's own PGS, genetic nurture. Conley et al (2015) do not find an effect, but Liu (2018) does. The educational system in the USA is widely different from that in Norway, and one can expect the confounding to be less severe in the USA than in social democratic welfare regimes due to the government-funded universally free school system. Further, Conley (2015) bases their PGS on an earlier iteration of the educational attainment GWAS (Rietveld et al. 2013), with the latest available (Lee et al. 2018) increasing the predictive power from 3 % to 15 %. Paper number 2 in this thesis do many of the same analyses done in Conley et al (2015), but with a larger sample, better prediction from the PGS as it's based on a later iteration of the educational attainment GWAS, a different educational outcome (achievement), and perhaps most importantly, in a different societal- and educational system context.

Gender may play a role in gene-environment interplay too, which was shown to be the case in the USA. Herd et al. (2019) show that the association between educational attainment PGS was stronger for men than for women for earlier born birth cohorts. However, as gender equalization took place over time, the difference became smaller over birth cohorts, as genetic influence rose for women. In the latest birth cohorts (1975-1982), there are no longer differences between men and women in the predictability from PGS on educational attainment. In other words, gendered constraints seem to inhibit women to realize the genetic propensities that the PGS for education captures. Twin studies also regularly find that heritability is higher for men when using birth cohorts from times when gender equalization was less advanced than in current-day Western societies (Branigan et al. 2013; Heath et al. 1985; Silventoinen et al. 2020).

For education, there is both a social background component and a genetic component. It's unclear whether the shared environmental influence causally springs from the family (Freese and Jao 2017), but it's clear that something that siblings share influences how much education they attain. Results from sibling correlations in Norway show that the sibling correlation for education is between 0,4-0,5 for same-sex siblings, and 0,3-0,4 for opposite-sex siblings (Wiborg and Hansen 2018). Judging by the high shared environmental

component in twin studies, it's plausible that a substantial portion of the sibling correlation is due to shared environmental influences for siblings.

For income, on the other hand, a different picture emerges. From Wiborg and Hansen's (2018) study which uses the same birth cohorts for education and income over time, same-sex sibling correlations are between 0,1-0,25, and between-sex sibling correlations a meager 0,01-0,15. Siblings are not at all that similar in income. Genetically informed studies may shed light on why this is the case. Hyytinen et al (2019) and Nielsen (2018) summarize earlier studies using pedigree designs and find that shared environmental influences typically range from 0-25 % of the variance explained, whilst heritability is 40-70 %. Mind you, many of the studies referred here are based on data from less egalitarian countries, where heritability can be expected to be lower than in countries with wide security nets like Norway. Compared to education, income is substantially more heritable, and less influenced by shared environments. Similarly, Benjamin et al. (2012) and Ørstavik et al. (2014) find that shared environmental influences are negligible, and genetic and nonshared environments dominate for income for a longer observational window of income measures (Benjamin et al. 2012) for recent birth cohorts of Norwegian men and women (Ørstavik et al. 2014).

The meager findings for sibling correlations could be due to income being highly heritable, being measured later in life than education, and being highly unpredictable for the individuals that seek to attain it. Siblings share on average 50 % of their genes. That leaves 50 % to be nonshared, and in a sibling model, this contributes to them being different, if genetics is predictive for income. Thus, the small coefficient for sibling correlations for income could reflect mainly genetic differences, and to a little degree shared environmental differences, and we could be left to expect sibling *differences* more than sibling similarities for income. The difference between education and income may come as a surprise to the general sociologist trained in homogenous social family background paradigms, but for many traits, heritability increase over the life course (Knopik et al. 2017). Education is a stable trait, you attain a level and never un-attain it over your life course. Income, on the other hand, could increase or decrease over time, even go negative, and have high fluctuations throughout the life course (Haider and Solon 2006). Income is less stable than education, and if it's under some causal influence from total family background, it seems to be genetically induced. Income is attained in a market-like setting, which is much more volatile than the educational system, characterized by being stable and predictable. Furthermore, the correlation between education and income is fairly low (Pekkarinen, Salvanes, and Sarvimäki 2017) in Norway due to the

compressed wage structure. The low sibling correlation in sibling studies in social sciences and the high nonshared environmental influence in genetically informed twin studies may simply reflect that income is highly influenced by randomness and luck (Freese and Jao 2017; Sauder 2020). I should point out that other economic outcomes may not be under similar influence from genetics. Wealth, for instance, is influenced by social background, but mostly through direct transfers of wealth (Black et al. 2020), and not through environments linked to childhood rearings, such as parent's education, household income, sibship size, or place of residence in childhood (Fagereng, Mogstad, and Rønning 2021). Studies using PGS confirm that income and other socio-economic traits are genetically influenced (Barth, Papageorge, and Thom 2020; Belsky et al. 2018; Hill et al. 2019; Hyeokmoon et al. 2020; Papageorge and Thom 2019).

I'm not aware of studies that investigate gene-environment interplay for income using twin methods in a similar vein we saw under the previously mentioned studies on education. The heritability of income is, however, higher for men than for women (Benjamin et al. 2012; Hyytinen et al. 2019; Ørstavik et al. 2014), indicating that gender interplay with how much genetics influence income.

Number of Siblings and Sociology

Nonshared environmental influences have proven elusive to the point where behavioral geneticists fear the “gloomy prospect” of environmental influences being idiosyncratic beyond scientific capability (Plomin and Daniels 1987). One of the few exceptions is birth order, closely linked to the number of siblings. The two are different, however, and would be so also in the behavioral genetic framework. The number of siblings is shared between all siblings in the same family, a shared environmental component if it's consequential for the trait at hand. Birth order is nonshared, as no non-twin siblings in the same family can share birth order. Nonshared environmental influences from twin studies thus indicate that nonshared influences like birth order matters as nonshared environments matter for twins, but as twins are born with the same birth order, one cannot use twin studies to investigate birth order differences. Instead, one must rely on other data sources and methods, which sociology has done to some degree.

A true landmark in revitalizing a broader understanding of social inequality issues in sociology, Blau and Duncan's seminal book include a chapter on kinship, including sibship size and birth order (Blau and Duncan 1967). They found that sibship size had an effect on attainment, whilst birth order did not. In the decades that followed, research into what structural causes led to differences between siblings was debated, culminating in the 2000s with decisive evidence that birth order differences were not confounded by sibship size. By and large, however, sociological explanation continued down the homogenous path of family environments influencing siblings in the same way even after Blau & Duncan's book. Is this a reasonable expectation? Like I've sketched up in the previous sections, nonshared environments in twin studies, the lack of predictive ability in general in social sciences, and low sibling correlations show that siblings can be expected to be dissimilar. Why that is, we know less of, but one of the few structural nonshared environmental influences that have been well documented is birth order.

Birth Order and Education

The three papers in this thesis on birth order all consider educational outcomes, and I focus mainly on that outcome here. Birth order does, not surprisingly, influence other traits that correlate with education too, like intelligence (Barclay 2014; Kristensen and Bjerkedal 2007; Sundet 2014) and income (J R Behrman and Taubman 1986; Bertoni and Brunello 2016; Kantarevic and Mechoulam 2006). For a long time, birth order differences were debated. Are they a reflection of one's birth order, or is there a spurious association due to large sibship sizes, which rise monotonically with birth order?

Sibship size is different between children from *different* families, whilst birth order is different between children *within* the same family (de Haan 2010). In other words, what matters the most, being born as number three, or having two siblings (Conley 2005:64–65)? Throughout the 1980 and -90s, a range of studies showed an association between sibship size and socio-economic outcomes (i.e. Blake 1981, 1986, 1989). Lacking adequate data and methods to identify what caused what, one of the main theories put forward by Blake was that number of siblings influenced the amount of resources siblings could access. A universal and parsimonious explanatory theory was formulated; *the resource dilution* theory. The more siblings you have, the fewer resources you can access from your parents, you have to share them in a zero-sum game with your siblings. The more siblings in the sibship, the fewer resources for you. In this theoretical framework, both sibship size and birth order could

influence your accessible amount. For a long time, sibship size had causal precedence in the theoretical landscape on why the number of siblings mattered (Steelman et al. 2002). In line with this, Rodgers (2001) claimed that the inadequacy of methods applied at the time led people to fall into the “birth order trap”; we all have a birth order and it’s easy to observe and talk about it, even though it’s not causal. Hundreds of previous studies, and several theories, were dismissed, due to what Rodgers believed to be a flawed methodological set-up.

A string of studies since the early 2000s has shown that birth order differences are real, and not confounded by sibship size or parental characteristics associated with having many children. Quite a few of them are based on Scandinavian data. Black et al (2005) investigate birth order effects in Norway on education measured in years at age 25. Using 650 000 families, the data are large enough to use family-level fixed-effects models with cluster-robust standard errors, models which take into account all things shared at the family level for each sibling. This means that the identification strategy employed can account for sibship size, as completed sibship size is a fixed entity for each sibship/parent. Doing the analysis separately by sibship size and considering birth order differences within each sibship size, Black et al show that there are differences between siblings according to birth order in all sibship sizes (2-10), at least when firstborn are compared to later-born siblings. The common pattern in birth order studies is that later-born siblings perform worse than earlier-born ones. Effectively, families are used as their own controls in the statistical identification to assure counterfactual assessment of birth order, we imagine what it will mean to be in birth order 2 compared to birth order 1 in sibship size 2-10, versus being number 3 in sibship size 2-10 and so forth. Family-fixed effects models have been the industry standard since the early 2000s, and a massive amount of research has shown that birth order matters (Barclay 2015; Barclay, Hällsten, and Myrskylä 2016; Black, Devereux, and Salvanes 2011; Bonesrønning and Massih 2011; Bratsberg and Rogeberg 2018; de Haan 2010; Kristensen and Bjerkedal 2010), and even that its effects may be intergenerational (Barclay, Lyngstad, and Conley 2021; Havari and Savegnago 2020).

Despite the overwhelming evidence of birth order differences, its causes remain more elusive. It’s not a “birth order trap”, but the birth order may have genetic or in-utero origins (Almond and Currie 2011; Foster and Archer 1979). Earlier-born children have a lower birth weight (Brenøe and Molitor 2017), and birth weight influence a range of socio-economic outcomes in adulthood (Conley and Bennett 2000), where lower birth weight on average leads to lower attainments. However, previous studies have indicated that birth order effects on

education and intelligence are environmental in origin, by controlling for birth weight (Brenøe and Molitor 2017), or using adoption (Barclay 2015) and early-stillbirths (Kristensen and Bjerkedal 2007) as quasi-experiments. Still, as we argue in Paper 5 in this thesis, there was reason to worry about potential confounding. We find that this is not the case, and show that birth order differences are indeed environmental in origin.

Being environmentally caused within families, birth order differences are an example of a clean social mechanism causing social inequality. Clean-cut social mechanisms are rare and offer a window into mechanisms creating social inequalities that could be operating elsewhere in society too. The two main theories on birth order differences emphasize resource dilution (Blake 1981, 1986, 1989), and differences in overall family environments and sibling interactions (Zajonc 1976, 1983; Zajonc and Markus 1975; Zajonc and Sulloway 2007). The vast majority of papers on birth order, however, does not have survey measures of parental resources and sibling interactions, making the leap from theory to empirical findings and back again mainly theoretical. Even if there is less risk for confounding from genetics for birth order differences, the causes of birth order differences within the family have proven hard to pin down. Comparing institutional settings, counterfactual assessment of meso- and macro-level institutions, just like with gene-environment interplay, may offer a promising avenue to learn more about what causes them.

Socio-Demographic Trends and Welfare State Norway

To provide the reader with at least a minimal background for the papers in this thesis, I briefly describe Norway as a case for social inequality studies, as I use data exclusively from this land.

Norway is one of the most equal countries in the world (UNDP 2018), whilst simultaneously having one of the highest GDP per capita in the world (Larsen 2001). Illustrating this, Norway saw the highest growth in real disposable income for households of all Western countries, with a 110% growth between 1979 and 2010. The percentage point growth was more or less equal across all income deciles, according to Our World in Data (2021). Social mobility is high (Pekkarinen et al. 2017), and Norway is a social-democratic welfare state (Esping-Andersen 1990, 2015). The extensive social policies introduced in the mid-1900s were explicitly put in place so that as many as possible could participate in the economic progress, regardless of the place of living, gender, age, social background, or

occupation (Thuesen, Thorsnæs, and Røvik 2019). Judging from genetically informed studies (Engzell and Troup 2019; Rimpfeld et al. 2018), we can expect genetics to have a larger influence in Norway compared to other less-equal countries, at least for education. However, education has shown to be considerably influenced by shared environments in Nordic countries too (Branigan et al. 2013), so I do expect that homogenous family environmental explanations are not fully confounded. For income, the evidence is less clear from genetically informed studies on what to expect as they are further in-between, but a higher heritability than for education is a reasonable expectation (Benjamin et al. 2012; Freese and Jao 2017; Hyttinen et al. 2019).

Norway ranks among the highest-ranking countries in gender equality (UNDP 2018). Norway was of course a part of the gender revolution that swept Western societies in the 1960s and 70s (England 2010). Moreover, the Norwegian welfare state explicitly aimed for gender equality to secure funding through dual-earner households and has fostered gender equality through extensive policies, such as equal-pay laws, institutionalized elder care and child care, making education free for all, and so forth. Higher gender equality can be expected to increase the influence from genetics, in particular, the influence from genetics for women can be expected to become more similar to that of men, at least for education (Herd et al. 2019). I'm not aware of any studies investigating gender, genetics, and lifetime income, but judging from Herd et al (2019), we can expect the same for lifetime income as for educational attainment.

Birth order can be expected to vary with macro-level institutional settings. Comparing less-developed countries to Western countries like Norway, birth order differences are in the opposite direction from Western countries in developing countries, later-born perform better than earlier-born (De Haan, Plug, and Rosero 2014; Tenikue and Verheyden 2010). In countries with abundant resources like Norway, birth order differences can be expected to be larger than in countries with many in poverty. Immigrants that arrive in Norway have the same universal welfare rights and access to education and work. Norwegian-born to immigrant children are well integrated into the Norwegian society, and some groups outperform Norwegian natives (Hermansen 2016). We employ this causal contrast in paper 3, where we investigate this paradox by looking at Norwegian-born children of immigrants, many of which have parents from non-Western countries that are far lower on the Human Development Index than Norway, and characterized by large inequality (UNDP 2018).

Norway is among the forerunners of the second demographic transition (Lesthaeghe 2010; Zaidi and Morgan 2017), characterized by sub-replacement fertility and diversity of union and family types. In many Western countries, immigration is keeping population numbers stable, or else they would dwindle due to natives having sub-replacement fertility (Anon 2017). We use two of these features of the second demographic transition developments to investigate social inequalities within families, Norwegian-born to immigrant parents and birth order, and multipartnered fertility and birth order. Being a forerunner, Norway offers a compelling case to investigate such differences, as other Western societies may be following suit. Knowledge on how social inequality comes about in light of new population dynamics is of pivotal interest for sociologists.

The Norwegian institutional setting being a welfare state undergoing the second demographic transition, being a dual-earner society, with extensive child care coverage and policies offer a golden opportunity to learn more about how sibling differences come about. I do so using a combination of counterfactually varying institutional and demographic settings. In the last paper, I incorporate genetics too, in a high-functioning sample with birth cohorts born in the 2000s, where most, if not all, the children probably have started in child care around age 1. Social settings like institutionalized child care may condition birth order effects and family environments more generally, and to our surprise we find that sibling interactions seem to matter more than previously thought in birth order studies. Not only do parental resources seem to matter less than previously thought, but sibling interactions may have arisen exactly because of institutions taking over from parents in child-rearing. This shows that expecting homogenous parental family environments may prove even more troubling in the coming years.

3. Methodology

Lundberg et al (2021:1) “make only one point: Every quantitative study must be able to answer the question: What is your estimand?”. Lundberg et al's program fit well with many of the perspectives I have already described in this thesis, and together they form a unifying thinking in science. In the 1920s, geneticist Sewall Wright invented the path analysis (Pearl and Mackenzie 2018; Visscher, Hill, and Wray 2008), which uses a closed system of nested relationships between variables in statistical analysis. Directed Acyclical Graphs (DAGs), which I will describe in more detail shortly, stem from this line of thinking (Pearl and Mackenzie 2018). Twin methods and the Structural Equation Modelling framework often used with them are also in direct antecedence to the genetic path models Wright used on mice. Barbara Burks (1928), a Wright contemporary, provided a pioneering, but overlooked due to gender discrimination, work on genetic confounding of social science phenomena. Although Burks was not recognized while she was alive, her legacy lives on. Blau and Duncan's resurrection of path analysis (1967) was inspired by Burks' early work (Pearl and Mackenzie 2018:309). Freese's counterfactual interpretation of genetics (Freese 2008) for sociology fits well with DAGs too, as DAGs are closely connected to counterfactual assessment of causality. Lundberg et al continue the resurrection of counterfactual causality and path analysis, providing sociologists with a framework that incorporates complex ideas in a more comprehensible way.

In this section, I aim to follow Lundberg et al's program for quantitative sociology, as it's been an eye-opener for (the shortcomings of) my work, some of which is already published as part of this thesis. I begin by introducing Lundberg et al's program for connecting statistical evidence to theory. Thereafter I place the papers in the framework before I go through each of the estimation strategies that are used in this thesis. Finally, I describe the data used in the papers.

Estimands: Connecting Statistical Evidence to Theory

An estimand is the target quantity, the purpose of statistical analysis. According to Lundberg et al, many quantitative sociological studies jump straight to data and regression coefficients when deciding what their goal of statistical analysis is. However, such a methodological approach becomes tautological, as the estimation object is defined within a statistical model

and cuts off links to other models that could accomplish the same goal. Figure 1 is adopted from Lundberg et al (2021) and shows the methodological recipe they propose.

Figure 1 - Estimands Framework

Theoretical estimands (TE) are defined by argument from theories and specified in unit-specific quantity (USQ) in a target population (TP). Consider Paper 1 in this thesis and let’s leave the gender and birth cohort aspect out for this example walkthrough. Some of the theory guiding paper 1 is partly described in its broadest form in the overall research questions for the thesis, and in the previous chapter: genetics may influence social inequality. I’ve argued that looking into genetics may provide insights into social inequality, as it’s unclear whether social background factors such as family environments or genetics cause sibling similarities. The goal is hitherto defined outside the statistical model, I use previous research to provide theoretical reasons to expect that genetics play a role.

The USQ clarifies the research goal, and in this case, it’s causal because I use a counterfactual intervention to investigate the theoretical estimand. The counterfactual part, a hypothetical intervention or potential outcome, is in this example linked to two things, the zygosity of the twins and the fact that the twins stem from different families. The USQ is the difference in shared genetic makeup between monozygotic twins (100 %) and dizygotic twins (50 %), and whether this difference is consequential for lifetime income. Another USQ, in this case, is whether differences in the shared environments between twin pairs from different families make the twins different in lifetime income. Further USQ that I have not exemplified for Paper 1 are nonshared environments, gender, and birth cohort.

The second part of setting the theoretical estimand is the target population, the set of units over which the USQ is aggregated over. The question remains whether the TP that is tractable is informative for theory in itself, or whether it's informative for the population of interest, i.e. the full population of Norway. I define the TP in question to be twins born in Norway over the 20th Century and claim that this TP is informative for the theoretical questions on what mechanisms cause social inequality between individuals in general.

Next, we need to link the TE to observable data and an empirical estimand (EE), a process which Lundberg et al (2021) label identification. In linking to observables, we can rely on Directed Acyclic Graphs (DAGs), which are closely linked to counterfactual causality (Pearl and Mackenzie 2018). In linking the TE to EE we use observable ones this time (no potential outcomes). The EE we choose (lifetime income, zygosity, and family origin) is informative of TE under key assumptions. Identification assumptions must be defended on conceptual grounds, and are themselves counterfactual. For each unit we observe, it is not possible to observe both conditions (i.e. being simultaneously MZ and DZ-twin). Thus, the empirical estimand is the difference in the observed outcomes relating to the difference in variance explained for correlations of MZ- and DZ-twins using twin identification strategies. Using DAGs formalize the linkage of TE to EE, and make it easier to make clear if there is a discrepancy between them, and states clearly the assumptions the chosen identification strategy makes.

Figure 2 is adopted from Hart, Little, and van Bergen (2021)², and is a DAG illustrating the identification problem in classical sociological accounts of social background paradigm if one considers the potential of confounding from genetics. The structure of Hart et al's DAG is somewhat similar to my DAG for paper number 2 in this thesis, albeit with different theoretical nodes (e.g. parental educational attainment vs parents' reading ability) and adding a node for child genetics.

² All figures from Hart et al's (2021) paper used in this thesis are adopted under CC-BY 4.0 licence

Figure 2 – Potential Confounding from Genetics

Shared genes between offspring and parents could in this DAG cause both child's reading ability and parents' reading ability, confounding the environmental effect of the number of books in the home, a variable often used to proxy social, cultural, and economic background. There are several problems with just using books in the home. First, as already illustrated, the proxy variable may be confounded by genetics. Second, the direction of causality could be the opposite (Engzell 2019), a child's reading ability may cause parents to accrue books for the child based on ability, which could be partly genetic as the DAG show. Blocking the paths from parents reading ability by controlling for it, for instance, would open a back-door path to genetics via the child, unbeknownst to researchers that are not theoretically concerned with genetics. This way, a DAG over the potential genetic confounding of our chosen EE for sociological TE (parental education/reading ability) shows us that the assumptions for the sociological identification strategy are that there is no genetic confounding at all, or that you don't care if it's caused by parental genetics or parental environments, else the theoretical model would be erroneous.

Lastly, in Lundberg et als framework, we learn from the data by doing estimation. The purpose of the estimation strategy is to estimate the EE. We use the estimation algorithm as a

tool to estimate components that are unknown to us. The EE in Paper 1 is identified with variance components from twin models. We have thus chosen twin models, and twin data to estimate the EE, which we purport to be informative for our TE.

In the following sections, I go on to discuss the estimation strategies and assumptions underlying them to inform the TE and EE in each paper. Table 1 shows how the papers in the framework I have just described, their TE, EE, DAG, and estimation strategy. I discuss the other TEs, EEs, DAGs, and estimation strategies in the next chapter³, where I summarize the contribution of each paper to the overall research questions. The DAGs in the table, are the most basic ones for the TE at hand. When I draw them as adjusted and unadjusted, it's to illustrate a *potential* identification problem. For instance, in paper 3, sibship size and parental characteristics are not adjusted for in the DAG, illustrating the potential confounding from parental characteristics causing some to have more children than others, and thus simultaneously causing children to have higher birth orders. In the paper, we use estimation strategies that cluster for shared characteristics of the parents, such as endogenous fertility decisions, solving the identification problem. The DAG thus shows the problem that needs addressing, that parents that choose to have many children may be less educated and provide different resources and family environments from those that choose to have many children. Of course, DAGs being conceptual and counterfactual, one could always add a U to signify unobserved confounders for each node in DAG if one has reasonable arguments for confounding. Here, I have left U out to simplify the graphical presentation.

³ DAGs are presented in larger format in the next chapter, should the reader suffer from a genetically induced condition, deteriorating eyesight, with environmental solutions not being adequate (glasses, lenses).

Table 1 - Research Design of Papers

Process:	Set the target: The theoretical estimand		Link to observables	Learn from data
Tools:	Unit-specific quantity	Target population	Identification	Estimation
Paper 1	Lifetime income rank person/twin pair i/f would realize if he/she/they had a) different zygosity b) different shared environment c) different gender d) different birth cohort	Twins born in 20th Century Norway		Standardized (constrained) twin model (ACE-model)
Paper 2	Standardized national test score that child i would realize if a) educational attainment polygenic score took a particular value, and b) parental educational attainment took a particular value	Children with European ancestry, born after 1999 in Norway		Ordinary Least Squares + PGS to adjust for genetic confounding
Paper 3	Lower secondary school point that child i would realize if it had a different a) birth order, and b) a different parental country of origin	Sibling children-of-immigrants in 11 ancestry groups		Ordinary Least Squares with fixed effects clustering at parents

Process:	Set the target: The theoretical estimand		Link to observables	Learn from data
Tools:	Unit-specific quantity	Target population	Identification	Estimation
Paper 4	Lower secondary school point that child i would realize if it had a different: a) full biological birth order, b) overall birth order, and c) sex of parent displaying multipartnered fertility	Siblings with parents displaying multipartnered fertility		Ordinary Least Squares with fixed effects clustering at parents
Paper 5	a) Educational attainment polygenic score; b) birth weight, and; c) educational achievement/attainment, that person i would realize if she/he had a different birth order	Siblings		Ordinary Least Squares with fixed effects clustering at parents

Twin Methods

Paper 1 uses twin methods for estimation of the influence of genetics and shared environments. Based on Structural Equation Models (SEM), ACE-twin models are themselves clear on what conceptual assumptions are made, they are in accordance with counterfactual causality, and can thus be tested empirically (i.e. see Appendix in Knopik et al. 2017). Figure 3 display the path logic behind twin studies (adopted from Hart et al. 2021). P stand for phenotype and T1 and T2 signifies twin 1 and 2. Our EE stems from estimating the paths in the model on lifetime income. The models informing EEs, analogous to the SEM framework, are only valid if the ACE model is a true reflection of reality (Knopik et al. 2017). In this model we assume explicitly that MZ-twins share 100 % of their genes ($r_{MZ} = [\text{correlation of}] 1.0$), and that DZ-twins share an average of 50 % of their genes ($r_{DZ} = [\text{correlation of}] 0.5$) in the model. Identical twins share as close as making no methodological difference 100 % of their genes (Jonsson et al. 2021), and siblings have on average 50 % similar genomes too (Visscher et al. 2006).

Figure 3 - The Classical Twin Model

Another explicit assumption concerns the environment, signified by the path between C – shared environment being correlated 1.0 for both sets of twins. It's assumed that twins have equal environments (the equal environments assumption – EEA for short). EEA is that MZ-twins are not treated more equally than DZ-twins. Note that this assumption would be invalidated only if MZ-twins are treated more equally *due to their higher similarity in itself* compared to other siblings. If MZ-twins have more equal outcomes because they are equally smart, compared with less equal outcomes for DZ-twins who differ in the innate ability and were assigned to different level of classes because of it (Conley and Fletcher 2017), we want the influence from the genetic similarity of MZ-twins to show for our trait of interest, even if mediated environmentally. The EEA has been criticized and questioned (Felson 2014; Horwitz et al. 2003; Richardson and Norgate 2005). Tests of the EEA, however, find that it generally holds, and if violated, is non-consequential for the general findings from twin studies (Conley, Rauscher, Dawes, Magnusson, and Siegal 2013; Conley and Fletcher 2017; Freese and Powell 2003).

An implicit assumption, that is an assumption that is not displayed in the figure, is that there is no assortative mating. Assortative mating is when spouses meet based on similar traits, which sociologists of course know occurs, like when highly educated people marry other highly educated people. However, here we are only concerned by genetic assortative mating, the concern that people that attain higher education are on average more genetically similar to each other than two random persons drawn from the population because something concerning their genes makes them more likely to pursue higher education. This could again lead to non-identical siblings being systematically more identical than 50 %. Genetic assortative mating occurs, but is much smaller than environmental assortative mating, at least in the US and the UK (Domingue et al. 2014; Hugh-Jones et al. 2016). Furthermore, if like begets like in spousal formation, a *positive* assortative mating, this would *underestimate* heritability in the ACE-model. If this is the case, sociological accounts are in even bigger trouble. Assortative mating, however, does not seem to confound results from twin studies in the US (Nielsen and Roos 2015), and the extent of assortative mating is typically lower in Nordic countries (Wagner, Boertien, and Gørtz 2020). Taken together, this assumption seems unlikely to be violated to a high degree in twin studies.

Lastly, ACE models assume that genetic effects are additive, i.e. that there are no interactions between genes known as epistasis. However, the vast majority of human traits are

additive (Hivert et al. 2021; Polderman et al. 2015), so the additive assumption is less problematic.

To sum up, twin methods rest on several assumptions which may be questioned. Be that as it may, a strength of twin studies is that they adhere to being very clear about the assumptions made, a virtue that cannot always be attributed to quantitative sociological studies. Morebeit, when the assumptions are tested, they generally hold (Conley, Rauscher, Dawes, Magnusson, and L. 2013; Conley and Fletcher 2017; Mönkediek et al. 2020; Scarr and Carter-Saltzman 1979).

Genome-Wide Association Studies and Polygenic Scores

In papers 2 and 5 we use Polygenic scores (PGSs). PGSs are based on Genome-Wide-Association Studies (GWAS), which are statistical brute-force bottoms-up studies of the human genome. Genomes consist of Deoxyribonucleic Acid (DNA) which comes in base pairs, of which genes consist. Genes are a conceptual description, genes are the basic unit of hereditary information. Genes contain code-encoding protein formation and ultimately the function of DNA, i.e. a gene encodes a message from DNA, to build a protein that makes a function in the organism. The human genome contains 20 687 genes, 1 796 more than worms, 25 000 fewer genes than rice and wheat, and have 3 095 677 412 base pairs (all numbers in this paragraph are from Mukherjee 2015). Most of the human genome does not encode for genes at all, and only a few of the base pairs vary among individuals. We are primates and our genomes are 96 % similar to that of bonobos and chimpanzees. However, those base pairs that do vary can have a profound impact, and mutations (copying-“errors” in the process of DNA duplicating itself via RNA) are often what cause still-births and several severe diseases. Thus, much of genetics, and in sociogenomics, in particular, is not about quantity in itself but the impact of genomic variation.

Single-Nucleotide-Polymorphisms (SNPs) are the most common type of genetic variation among individuals. A SNP is a variation in a single base pair of DNA, which happens at about every thousandth base pair, where one person will have a C and another might have a T in their genomes. Still, humans are 99 % genetically identical in their SNPs, but the small fraction of genetic dissimilarity is what we are studying (Mills, Barban, and Tropf 2020). GWASs regress the trait of interest, such as educational attainment, on each SNP and use Bonferroni-correction to see which SNPs reach significance levels using p-value

thresholds adjusted to base-pair sample size. Most SNPs have no known influence on human variation, but this is a field of rapid development. Summary statistics from GWAS, which summarize which SNPs were found to be associated with the outcome for the sample at hand, can be used to create a summary score of all SNPs for prediction on other peoples genomes. A polygenic score summarizes all the genetic variations captured in a GWAS into one value for each person. Figure 4 (Hart et al. 2021) shows part of the process of creating a PGS.

Figure 4 - Genome-Wide-Association Studies and Polygenic Scores

Polygenic scores also signify a new conceptual direction in behavioral genetics and sociogenomics. Poly stands for many, as in many different genetic variants influencing an outcome, and is so widespread that a fourth law (cf. Turkheimers (2000) three laws) of behavioral genetics has been defined: “A typical human behavioral trait is associated with very many genetic variants, each of which accounts for a very small percentage of the behavioral variability.” (Chabris et al. 2015). Crucially, PGSs are predictive out-of-sample, indicating that they capture actual genetic signals. Another advantage of the PGS is that peoples genomes do not change (yet), PGS capture genetic variants that are fixed at birth and are intrinsically prospective (Freese 2018). We can therefore use PGSs as “molecular tracers” following developmental processes over a life course (Harden et al. 2020). Moreover, PGSs

can be used in much smaller sample sizes than the original GWAS and retain predictive power, as one can use all available SNPs, not just the ones that reach genome-wide-significant after the Bonferroni-correction (Mills et al. 2020). Combined with a rapid increase in social science samples with genotypes available, PGSs are bound for widespread inclusion in the social sciences over the coming years, the sociogenomic-era is here (Conley and Fletcher 2017; Freese 2018; K Paige Harden 2021; Harden and Koellinger 2020; Mills and Tropf 2020). However, being an active field of methodological and conceptual development, in the last three years or so questions have been asked whether PGSs pick up solely genetic signals.

Population Stratification, Family Trios, and Genetic Nurture

PGSs capture signals from genetic variants in a person's genome, but may pick up environmental variation too. Environmental variation in PGSs may stem from assortative mating, population stratification, or dynastic effects (Morris et al. 2020). If PGSs pick up environmental signals, they can inflate the true values of population estimates, and violate the counterfactual aspect of using PGSs to investigate influence from genetics on social phenomena. I described assortative mating as a potential problem for twin studies, and the problem is the same with PGS; if individuals that pursue higher education are more similar in their genomes on average than two random persons from the population, and consequently perform successful coitus viewed from the evolutionary perspective (i.e. the formation of gametes follow and in uterus conditions sustain the gametes 44 weeks or so), this would lead to biased estimates of causal effects of the genotype on education in the offspring generation (Morris et al. 2020). However, when this is tested for molecular genetic data, assortative mating is non-consequential for heritability estimates (Conley et al. 2014).

Population stratification concerns population structure and differences in social outcomes over time. A much-used example is a chopstick-gene example (Hamer et al. 2000). If you scan people of the world and look for genes for eating with chopsticks, you will find them. But there are obviously no genes for eating with chopsticks, just like there are no genes for eating lutefisk. An adopted child from Korea being raised in Norway will not eat with chopsticks. Rather, the chopstick-gene example shows that it's the results of ancestry differences due to allele frequencies varying between sub-groups of the human species, as our migrational history spreading successfully throughout the globe is fairly recent from an evolutionary perspective.

It's believed that every single human individual on earth shares a common ancestor, a female that lived in the Rift Valley in Africa around 200 000 years ago, known as Mitochondrial Eve (Mukherjee 2015). Tracing our migrational origins through DNA-sequencing (Li et al. 2008) geneticists confirm the “recent African origin of modern humans” theory. As humans are thought to have migrated (and stayed) out of Africa less than 100 000 years ago, and only a few people migrate to each subsequent region of the world as we continued to spread out on foot, the result is a step-wise populational bottleneck. Consequently, the genetic variation in human genomes is highest in Africa. The farther from Africa, the lower the genetic variation due to populational bottlenecks (Conley and Fletcher 2017). Astonishingly, individuals from tribes living a few hundred kilometers apart in sub-Saharan Africa have larger DNA differences than any two people of European or non-African ancestry (Schuster et al. 2010). Following our migrational patterns throughout the world are non-causal SNPs that happened to be residing with their bearer. Han-Chinese have somewhat different SNPs than European ancestry Whites. And as Han-Chinese on average eat more with chopsticks than European ancestry Whites do, a GWAS might find SNPs seemingly associated with eating with chopsticks should it not adjust for such spuriousness.

To alleviate population stratification, researchers may apply principal component analysis (PCA) (Mills et al. 2020). Principal components (PCs) reduce the data dimensionality from high to low by identifying axes that contain the greatest genetic differentiation among individuals. We can then use predicted scores on PCs as controls, as they capture some of the differences between subpopulations in their alleles. A common standard is to control for 10-20 PCs when using PGSs between unrelated individuals to control for population stratification. However, even controlling for 100 PCs does not remove all population structure, even in a homogenous population such as those of European ancestry living in the UK (Haworth et al. 2019). Increasing attention to whether PGSs pick up environmental variation has led to studies showing that peers genomes may influence the outcome of an individual independent of the person's genome (Domingue et al. 2018), which would bias the results from genotype-phenotype associations. A recent development has been to consider whether parents' genomes influence their own children's outcome over and above what children's genomes do (Conley et al. 2015), later denoted genetic nurture by geneticists (Bates et al. 2018; Kong et al. 2018).

The current gold standard to use family data with genotypes for both parents and offspring, family “trios” (Harden and Koellinger 2020). By using trios one can get causal

estimates of the child's genotype on the outcome. Besides one can use the parental genotypes as predictors for social background phenomena, like home environments that sociologists usually invoke for their explanations. In this model, social background effects are free from genetic confounding too, as both children's and parents' genomes are already in the model. Any effect of parental genomes has to be environmentally mediated, as there is no genetic pathway as children's genomes are already in the model. Some suggest we think of genetic nurture as an “adoptive parent” (Harden and Koellinger 2020). Figure 5 (Hart et al. 2021) shows what genetic nurture designs look like graphically.

Figure 5 - Genetic Nurture

In paper 2, we use rich trio data where we include parents' genomes in the analysis to estimate our empirical estimand. This “gold standard” approach allows us to look past some potential problems (e.g. population stratification confounding the effect of child's genetics on the child's outcome).

Family-Level Fixed Effects and Birth Order

In papers 3-5, we use family-level fixed effects with standard-robust errors clustered at the parental level. This estimation strategy is necessary to distinguish between sibship size (shared among all siblings in a sibship, i.e. shared environment) and the nonshared environment that birth order is. If we look at the DAG drawn in Table 1, it's clear that sibship size and birth order could be confounded by parental characteristics, which was the methodological critique of earlier birth order studies made by Rodgers (2001). One could use control variables to try to remove between-family differences and get closer to birth order estimates free from parental or sibship size confounding. However, it's impossible to control away all possible confounders, and the DAGs would be very complex and one could end up controlling for mediating variables. Like genetic studies where the gold standard for causal inference of PGS free from environmental confounding is family data, either comparing siblings or family trios (Belsky et al. 2018; Harden and Koellinger 2020; Mills and Tropf 2020), family data is the gold standard of birth order studies too.

By clustering on the family level (parents), we are in a sense clustering the shared environment of the siblings, and considering only the variation that is different between siblings in the same family according to birth order. This estimation strategy partitions out family-level influences that are constant over time, such as similar rearing practices, socioeconomic status, and importantly sibship status. Our estimation strategy is using ordinary least squares and uses dummies within each family to contrast the effect of being later-born to being firstborn. Simulations show that birth order effects can be causally estimated using this strategy (Kravdal 2019).

Although the family-fixed effects clustering have been hailed as the gold standard for estimating birth order effects, there may be problems with confounding also in this estimation strategy. Birth order differences could be caused by genetic or prenatal environments, which would indicate genetic confounding in the first example, and less-clear cut environmental pathways in the second example. In paper number 5, we examine this potential confounding by looking at PGS and birth order, birth weight and birth order, and lastly controlling for them in a family fixed effects set-up. Results from previous studies looking at potential confounding from genetics and in-utero environments that find that birth order differences are not confounded (Barclay 2015; Kristensen and Bjerkedal 2007; Muslimova et al. 2020) hold.

We thus show that the family-level fixed effects strategy for birth order studies is valid, and a side finding is that studies comparing siblings' PGSs are not confounded either.

Data Sources

The empirical estimands in this thesis are all based on administrative data sources from Norway. Administrative data sources are very valuable for several reasons (see i.e. Figlio, Karbownik, and Salvanes 2016; Røed and Raaum 2003). Signified by their name, these data are collected for non-scientific purposes. Scandinavian countries like Norway have been collecting data on their citizens for decades, to give public entities data and statistics to make informed decisions. All Norwegian citizens, and persons who obtain work in Norway, are given a personal identification number (ID) to link between the various registries that inform public entities. With the right permissions, one can use administrative data for research. The data used in this thesis are obtained through research projects with permission from the relevant ethics bodies. Data for papers 1, 2, and 5 are obtained through the European Research Council-funded project *OPENFLUX*⁴, spearheaded by PI Torkild Lyngstad. Data for paper 4 and 5 are obtained through the project *Ethnic segregation in schools and neighborhoods: Consequences and dynamics*⁵ financed by the Norwegian Research Council., The PI for this now-ended project was Gunn Birkelund. Given that all papers are based on data obtained via projects which have gained permission through ethical boards and data authorities, I do not go at length describing ethics related to privacy, data collection, and processing here.

Due to stringent security- and data protection requirements the different projects all obtain their datasets through Statistics Norway with non-identifiable ID numbers, and are kept on different servers. The underlying administrative sources, however, are the same. Working on these datasets usually involves the same initial set-up for each project. It's common to start with the Central Population Registry (CPR), which contains basic demographic information on the full population of Norway, and I use CPR in all my papers. The CPR allows for kinship linkages, which is pivotal in every paper. The CPR allows us to identify the parental country of origin in paper 3, whether parents have children with multiple parents in paper 4, and additional twins in paper 1, to mention just a few of its uses. Having made the CPR-based data file, one can link this dataset to data from other sources, such as the National Educational database (NuDB), which I use for four of the papers in this thesis. The NuDB contains many

⁴ See: <https://www.sv.uio.no/iss/english/research/projects/openflux/> | <https://torkildl.github.io/openflux/>

⁵ See: <https://www.sv.uio.no/iss/english/research/projects/segregation/>

variables collected for administrative purposes, and for adults (paper 2, 3, 4, 5) I use mainly the NUS-variable, which is very similar to the international standard for education ISCED (Barrabés and Østli 2017). Another variable from the NuDB used in two of my papers is lower secondary school points (papers 3 and 4), which is the sum of all grades attained in lower secondary school that is the basis for further enrollment to high school. Lastly, I use data on standardized tests performed in primary and lower secondary school from the same registry.

In paper 1, we use the CPR and data from the twin registry of Norway. We base the main analysis on the twin registry and supply with data from CPR on the full population. The twin registry is not based on administrative sources, it's based on several surveys of twins performed over time and thus based on consent (Harris, Magnus, and Tambs 2006; Nilsen et al. 2012, 2013; Tambs et al. 2010). It contains zygosity information and covers many birth cohorts born in the 20th Century. We link these two sources to administrative data on income from tax reports, covering the years 1967-2016 for the full population of Norway. This allows us to investigate lifetime income rank with twin models as we have very many data points to use for our estimate, and use birth cohorts and gender as additional causal contrasts.

Data from administrative sources are very good to circumvent common support problems, which can be a concern for estimands using observational data (Lundberg et al. 2021). Common support problems arise when the data at hand does not contain the observations necessary for your chosen theoretical estimand. For instance, papers 3 and 4s are the first of their kind by considering two different sub-groups of the population and birth order differences, immigrant families and multipartnered fertility parents. Although researchers around the world are free to imagine, following Judea Pearl's terminology for counterfactual causality, how birth order differences could play out in such settings, we're not aware of anyone who has done so previously to our studies. The reason is probably related to lack of common support.

Researchers considering birth order differences are mainly quantitatively oriented and do not have support in the data they have access to investigate theoretical estimands as we do in paper 3 or 4. To reach adequate statistical power, one would need around $N = 800$ in each subgroup of the population to identify birth order differences of similar magnitude to what other studies find (Kristensen and Bjerkedal 2010). Even though Norway has a limited population size compared to many other countries, we do have data on the full population far larger than many other data sources from more populous countries. Most survey data does not

contain many thousand individuals to allow for sub-group analysis as we do, we have good common support for our chosen TE. For paper 1, on the other hand, common support could be a problem. If I wanted to estimate twin models to investigate gene-environment interactions like heritability at different levels of socio-economic status (Baier and Lang 2019; Figlio et al. 2017; Scarr and Weinberg 1978), I would need observations over the distribution. As the rate of twinning is around 10 in 1 000 deliveries (Monden, Pison, and Smits 2021), I would run into a lack of common support, for instance in considering the heritability of income in children with very wealthy parents (top 1 % of the distribution).

Administrative data allow for the target population to be quite close to the actual population. No Norwegian has consented to be subjected to analysis through administrative data. Rather, it's decided by several official bodies whether your research is permitted to use the data (i.e. Datatilsynet, Regional etisk Komite, Statistics Norway). As the sources are not dependent on consent and are from administrative sources, which are often reported directly from one system to another (e.g. hospitals to medical birth registry; income from employer to tax authorities; grades from schools to national educational database), they are of very high quality. They suffer little attrition, no non-response, and little if any measurement error. In sum, administrative register data have many comparative advantages. However, administrative data sources have their limitation too. They do not measure attitudes, actions, beliefs, or subjective experiences. There is a gap between the observations the researcher does in her data and the meaning, reasons, and subjective conditions that guided the individuals' observed result at any given time. Nonetheless, administrative data sources allow for studies that would be impossible without such broad and deep data.

In papers 2 and 5 we use the administrative data described above, and molecular genetic data which are collected with consent through the Norwegian Mother, Father, and Child Cohort Study (MoBA) (Magnus et al. 2006, 2016; Rønningen et al. 2006). MoBA is a population-based pregnancy cohort administered by the Norwegian Institute of Public Health. Adhering to the “gold standard” of having genotypes of both parents and children, it's one of the largest of its kind, if not the largest. Combined with the administrative data, it's truly a unique source for sociogenomics and many other scientific fields. All pregnancies in Norway between 1999-2008 were sampled, and 41 % of the mothers consented to participate. MoBA consists of 114 500 children, 95 200 mothers, and 75 200 fathers. However, all are yet to be genotyped, as there is a genotype and quality control queue. We currently have genotypes of around 98 000 individuals. We perform genomic quality control ourselves and construct PGSs

for all members of each family following guidelines of best practice (Mills et al. 2020). The MoBA also contains many survey waves. We use data from one of these where the mother report birth weight and birth length, which we use to investigate birth order differences. The accuracy of self-reported birth weight and birth length could be questioned. However, I see no reason why measurement error should vary systematically between birth orders. Self-reported errors would only be a problem if they could cause systematic estimation errors. Like virtually all other surveys, MoBA is a high-functioning sample where there is a correlation between participation and educational attainment, reducing the external validity. Similarly, external validity is reduced as a larger and larger part of the Norwegian population has non-European ancestry. Currently, genomic data is very homogenous with an under-representation of non-European ancestry genomes in GWAS', meaning that it's not possible to analyze non-Europeans in studies like the one I conduct. This is "solved" by removing individuals with non-European ancestry.

4. Summaries of the Five Papers

In this thesis I investigate two overall research questions: 1) What can we learn from including genetics in studies on social inequality 2) What can we learn from including nonshared environmental influences like birth order in studies on social inequality? The first overall research question is related to the potential genetic confounding of family environments in sociological inquiries of social inequality and mobility. I address this in the two first papers in the thesis. Are genetic and social influences independent of each other, or are our sociological accounts confounded? In this section, I include the DAGs from Table 1 to clarify the basic identification contribution each paper makes.

Paper 1

This piece of work investigates whether a long-standing finding in social inequality research is robust to genetic confounding, by looking at lifetime income. With comprehensive income data from administrative sources and a large twin sample ranging many birth cohorts, we investigate whether the influence from genetic, shared environmental, and nonshared environmental influences changed as Norway changed over the 20th Century, with particular emphasis on changes for women. The DAG shows the problem of confounding, and how the paper also speaks to overall research question number two in this thesis, by estimating how large environmental sibling dissimilarity is.

Figure 6 - Directed Acyclical Graph - Paper 1

The theoretical estimand is linked to whether the unit-specific quantity of various genetic similarity (MZ vs DZ-twins) cause differences in lifetime income rank. Similarly, shared environments are causally contrasted in the unit-specific quantity, do they cause differences in lifetime income rank? Lastly, nonshared environmental influences are also contrasted by differences in outcomes between MZ-twins.

At first glance, shared environmental influences are very high for many birth cohorts, higher than genetic influences, which is a very unusual finding using behavioral genetic methods (cf. Turkheimers second law). Seemingly, shared environments, encompassing sociological family background, are highly influential until the last birth cohorts, confirming that family background could have causal precedence, and may not be under great risk of confounding. However, upon closer inspection, this is due to changes related to gender equalization, which can be attributed to the macro-level theoretically, not causal influences from middle-range family environments per se. Considering lifetime income rank separately by gender, shared environmental influences are negligible in most cohorts. Instead, genetic

and nonshared environmental influences dominate. This shows that sociological accounts highlighting homogenous family environments as causal forces for income is therefore considerably confounded, but not fully. Furthermore, this paper also has a nonshared environmental component finding on research aim two, sibling dissimilarity. Nonshared environments are highly influential for lifetime income attainment. Income is something most individuals in a society care about and is perhaps the most sought-after good in modern societies. This shows that even a highly sought-after good has a high level of idiosyncrasy, highlighting the elusiveness that can be expected in social science prediction, and that it's problematic to attribute influence to family environments without considering genetic influences. As such, this paper points to new theoretical questions for sociology, i.e. should we be giving family environments causal precedence in our theories?

Paper 2

In *Social and Genetic Effects on Educational Performance in Early Adolescence*, we look into intergenerational transmission of education. Using molecular genetic data with genotyped children and parents, we assess the intergenerational transmission of genetic and social benefits for scores on standardized tests in adolescence. Contrary to twin methods, molecular genetic studies do not yield ACE-estimates allowing for variance dispersion into genetic-, shared environmental-, and nonshared environmental influences. However, they offer a different methodological approach to investigating whether shared family environments are confounded or not, by incorporated molecular genetic measures in models where parental measures are included. The TEs are similar to paper 1, would a different unit-specific quantity (PGS / parental educational attainment) over European ancestry children born after 1999 cause differences in national test scores in adolescence? Further, we consider the TE that concerns confounding.

Figure 7 - Directed Acyclical Graph - Paper 2

Recall the most basic DAG from Table 1 that guides this paper. Classical sociological studies are prone to confounding from genetics as parental education, and the family environments sociologist often claim are causal influences on child's education, could be confounded by parental genetics causing both children's genetics and causing parents to attain higher education themselves. The association could thus be spurious, partly confounded, or not confounded at all.

In this paper, we use regression to assess to which extent intergenerational transmission of social background effects like parental educational attainment is confounded by genetics by including variables for the unadjusted nodes in the DAG. For education, children's genetics and parental educational attainments both play a role. The combined effect of mothers and fathers measured educational attainment is larger than the child's educational attainment-linked genotypic variants. However, the child's genotype is still the largest single-measure predictor. The confounding of the classical sociological family background effect as measured by parental educational attainment is found to be only around 1/5th. In other words,

genetics and social background both play a role in test scores, but they are largely independent of each other. This finding could indicate that part of the nonshared “puzzle” is a genetic effect, at least for non-twin siblings, as the PGS is highly predictive, but not at the expense of parental educational attainment. Looking further at whether genetics confound the social background effect, we consider whether there is an effect of parental genes, known as “genetic nurture”. This would indicate confounding also at the parental level as the effects of parents’ attained educations would be confounded by (parents’) genetics. We find a genetic nurturing effect, showing that part of the intergenerational effect of parental educational attainment is influenced by parental genetics. The paper shows that for education sociologists should have less of a worry of confounding. Results are in concordance with studies based on pedigree methodologies, as education is one of the few outcomes for which twin studies find that shared environments play a large role. However, genetics and shared environments are of around equal strength in pedigree studies, and our predictions from PGSs are lower, as PGSs’ are noisy measurements. The predictions from molecular genetic data and techniques may be improved in the near future, with within-family GWAS’ (Howe et al. 2021) and whole-genome sequencing may improve predictions. If this turns out to be the case, confounding may be more severe than this study finds.

Moving on to research aim two in this thesis, I consider what sociologists may learn from nonshared environmental influences, which are understudied in sociology compared to shared environmental influences. Besides paper 1 that speaks to the overall nonshared environmental influence on income, I have three papers on birth order (paper 3-5), which is one of few structural nonshared environmental influences. They have different twists to shed light on the phenomena.

Paper 3

The first TE for this paper is whether birth order causes differences between siblings. We use the industry-standard family fixed effects with clustering on parents and dummies for each birth order after firstborn. Clustering on parents holds parental characteristics that vary between sibships constant, such as potential differences between parents who choose to have more children. The “threat” to identification presented in the DAG is therefore remedied by the estimation strategy. We first establish that there are differences between siblings in the same family in lower secondary grades for Norwegian natives. Similar to what other research

has found, earlier-born siblings, perform better than later-born siblings in the same family when the Norwegian natives are considered.

Figure 8 - Directed Acyclical Graph - Paper 3

Having investigated this TE, we investigate birth order differences with variation at the parental level according to country of origin. In other words, we are varying at the level of parental characteristics in the DAG, whilst using the same identification and estimation strategy, plus some additional statistical tests to allow interactions between the parental country of origin to be freely estimated. For sibships in some parental origins, there are no birth order differences. This is surprising, given that birth order theories are universal, and one would expect no variation according to parental origin if this was the case. Still, in many immigrant families, the birth order effect is very similar to that of Norwegian natives showing the strength of this within-family inequality mechanism. We believe our analysis is informative for theory, and argue that birth order effects may vary between immigrant families with a different parental country of origin because of differences in parenting

strategies, child care attendance, and peer influence. As birth order differences are one of the few social inequality mechanisms that are free from genetic confounding and happen within the family, they offer one of the clearest cut social inequality generating mechanisms. It may give us pointers to how social inequality comes about more generally.

Paper 4

In this paper, we take a different approach to the nonshared environmental birth order question. We use the same data sources as in paper 3, allowing us to investigate subgroups of the population based on fertility decisions. More and more people procreate with multiple partners, they have children with more than one person. This fact allows for an exploration of birth order effects in that context. As older children who have younger half-siblings will usually have experienced parents breaking up which may influence children's educational performance, the answer to whether there are birth order effects in this context is not obvious. We use the same identification strategy as in the previous paper, family-level fixed effects with cluster robust standard errors and dummies for birth order, comparing firstborn to later-born.

Figure 9 - Directed Acyclical Graph - Paper 4

The first TE is whether there are birth order differences in families where parents have children with multiple partners. There is. In the next TE, we drill deeper into the causal contrasts of full biological siblings versus half-siblings. We find birth order effects on lower secondary grades in full biological and overall (including all full- and half-siblings) sibships, once more showing the strength of this nonshared environmental within-family inequality generating mechanism. In the last TE, we consider whether biological- and overall birth order effects vary according to parental sex. Here, multipartnered mothers yield birth order differences across their sibships and multipartnered fathers within sibships only. This allows us to poke at the mechanisms generating birth order effects, informing the overall theory on birth order differences. We attribute this to residential arrangements following union dissolution, which fits with resource dilution theory. However, we also emphasize the theoretical possibility that maternal cultural resources may be more important than paternal. We raise the possibility that mothers may be more important for birth order effects in education, and thus in child development more generally. This fits well with the findings from genetic nurturing effects, where maternal genetic resources seem to be the ones that make the

differences (Kong et al. 2018, but see Paper 5 in this thesis for further complications to this argument). Again, as our methods allow us to alleviate concerns of confounding, and happen within-family, this offers a window into how inequality in educational outcomes is generated more generally.

Paper 5

In the last paper, we investigate whether birth order effects themselves may be confounded by genetics, using molecular genetic data and survey data, in combination with administrative data covering the full population. The first TE in this paper is linked to the unit-specific quantity of educational attainment polygenic score. We investigate whether the PGS vary according to birth order, which would confound many research designs in a plethora of sciences that use sibling comparisons to secure causal estimates.

Figure 10 – Directed Acyclical Graph - Paper 5

The DAG shows what is the threat to identification in birth order studies, including paper 3 and paper 4 in this thesis. If there are genetic differences linked to education that vary with birth order, the theoretical explanations in birth order studies would not be correct. Our EE shows that it does not vary with birth order. The next TE is linked to in utero-influences, which would indicate different mechanisms than those that are usually used in birth order studies. We look at differences in birth weight and birth length according to birth order. Here, our EE show that there is a *positive* birth order effect, later-born are heavier than earlier-born. As lower birth weight is associated with lower socioeconomic attainments, this is in the opposite direction of negative birth order effects. Thereafter, we investigate whether including genetic and birth weight in a model where national test scores are the outcome influences the

birth order differences, which they do not. We conclude that birth order effects are not confounded by either genetics or in utero differences as measured by birth weight, and are social in origin. Having already informed several longstanding theoretical claims by showing that birth order differences are caused environmentally post-natally, we then go on to see whether we can figure out what social mechanisms cause them. We look into the potential influence of maternal age, birth spacing, differences according to parental educational attainment, gene-environment interactions, and maternal-, paternal- and sibling genetic nurture effects. A surprising find in one of our supplementary analyses is that “genetic nurture” effects may be sibling interactional effects, furthering the depth of the problem for sociological accounts of homogenous family environments causing similarity. Again, this paper shows the strength of the social birth order effect, confirming it is fully social in origin and shows that the expectation that family environments are homogenous is ill-fraught.

5. Discussion

How did thou inherit your socio-economic status? If you ask this question to a sociologist concerned with social inequality, it's a good bet that the answer would involve parents. However, what parents do to influence their offspring is less clear, even if the general sociologist point to parental socio-economic attainment having intergenerational effects. Throughout this thesis, I hope to have shown that answering this question with (implicit) arguments that parental attainment work through the family environment and cause siblings to be similar with regards to socioeconomic outcomes may be problematic. In the first overall research question, I asked what we can learn from including genetics in studies of social inequality. In the second overall research question, I asked what we can learn from including nonshared environmental influences like birth order in studies of social inequality. I discuss the contribution of each paper to our understanding of what causes inequality; family environments, genetics, or nonshared environments.

Sibling correlations are often used to assess the impact of all influences that are shared among sibling children in a family. However, sibling correlations are surprisingly small, only 0,01 – 0,15 for income for different-sex siblings in Norway (Wiborg and Hansen 2018). The sibling correlation for education for the same birth cohorts is 0,3 - 0,4. Sibling correlations may be caused by siblings being genetically similar, as they share 50 % of their genes. However, genetics may also contribute to differences between siblings, as 50 % are *not* shared between them. Paper 1 in this thesis shows that, when looking past the gender aspect and considering lifetime income rank within sex, sibling similarities in income are by and large genetic in origin. This shows it's highly problematic to assume that similarities in family environments chiefly cause siblings' similarity.

A further implication comes the finding for the nonshared environment for sex-specific lifetime income. These are as high as genetic influences, which show that we can expect twin siblings to be influenced by environments which they do not share. If twin siblings are that different in lifetime income, we should not expect non-twin siblings to be any less different, unless sociologists suddenly begin to emphasize niche formation theory (Conley 2005), i.e. that siblings actively try to be different from each other forming "niches". Sibling dissimilarity is to be expected, and we should expect that a large chunk of any similarity is genetic in origin. Accordingly, empirical studies calling on theories that emphasize family environments bear an evidentiary burden that this is indeed the case. With

Lundberg et al (2021), we can ask “which estimand shows that family environments cause sibling similarity?”.

Sibling correlations may mask big underlying changes in causal influences from genetics as measured by A and shared environments as measured by C in twin models. This could indicate that sociological models without a genetically informed design are missing vital information on what causes inequality, and it may be even more structured than we currently believe (Branigan et al. 2013). Using twin studies as “Gini-coefficients” (K Paige Harden 2021), they may be more suited to speak to the equality of opportunity than studies that are uninformed by genetics, in which sociologists often make a causal leap from statistical “chances” to equality of opportunity (Swift 2004).

In paper 1, I show that one can use twin studies to assess the measure the functioning of the social and genetic “lottery” (K Paige Harden 2021) over time in Norway, with particular emphasis on gender. It should be noted that the trend towards equalization between genders in lifetime income as measured by rising heritability over birth cohorts speaks nothing to whether this is the result of benign processes or not, as I have not pinpointed the actual mechanisms causing it. For example, in an educational system where college- and university admission could be influenced by heritable physical traits that increase your competitiveness in varsity sports, like height, the heritability of college- and university admission would be partially influenced by the heritability of height. However, few would contest that this mechanism is indicative of achievement versus ascription (Freese 2008; Freese and Shostak 2009). Nonetheless, most would agree that high shared environmental influence which could stem from differences in parental income, rearing environments, parental educational attainment, and schools, would not indicate equality of opportunity. As such, would we not want heritability to be something, as an indication that at least it’s not solely shared environments that influence life chances? Or would we want it to be all nonshared environments, i.e. entirely random? It’s beyond the scope, and ability of the author, of this thesis to even attempt to answer such normative philosophical questions, but my results raise interesting related questions for the able mind.

In the second paper in the thesis, we consider the intergenerational social and genetic transmission of education and use a molecular genetic design. We are considering achievement at an early age (grades 5, 8, and 9), where parental influence can be expected to be higher than at later ages. Here, we show that genetics, as measured by PGSs, predict of equal strength to parental educational attainment for national test scores. However, we also

show that the genetic confounding of the social effect of parental educational attainment is low. Taken at face value, this furthers the point I made in the previous paragraph; perhaps inequality is even more structured than we believe? One key takeaway is that sociologists should worry less about confounding of their analysis, as the two predict independently of each other. This is in accordance with findings from twin studies, where education is one of the few traits that show high shared environmental influence. Furthermore, it's in line with higher sibling correlation in education compared to income (Wiborg and Hansen 2018).

Education and income seem to be variously influenced by shared environments, with education being one of the few human traits that are highly influenced by shared environments (Branigan et al. 2013; Silventoinen et al. 2020), and income hardly being influenced at all (Benjamin et al. 2012; Hyytinen et al. 2019; Ørstavik et al. 2014). There could be several reasons for this. Education happens in a highly predictable system, and early achievement begets later admission. Early achievement happens whilst the achiever resides with the parents, making it easier for parents to influence both achievement and thus later attainment, as well as values. Income, per contra, is attained mainly in adulthood, when the attainer typically does not reside with her parents, at least in most Western countries. When you are “out of the nest”, your peers are no longer primarily your parents, but your friends, colleagues, and networks. Influences from any one of these would be attributable to nonshared environmental influences in the twin methodology, insofar as you do not share them with your sibling.

Likewise, luck may play a large role in occupational trajectories and income attainments (Figlio et al. 2017; Jencks et al. 1972). Some wonder whether we should view genetic endowments as results of luck too (forthcoming K. Paige Harden 2021; Hyeokmoon et al. 2020; Sauder 2020). The world and our societies are highly unpredictable and influenced by chance, randomness, and luck to a much larger degree than we like to admit. Social scientists can't even predict a highly structured trait like education with very rich family data (Salganik et al. 2020). Should we expect the general Jane and Joe to do better for income? When this was written, in April 2021, humans all around the world are having a hard time predicting what will happen the next month, as we're experiencing the greatest pandemic since the Spanish flu 100 years ago (Christakis 2020). The structures that guide agency for income are less predictable than for education, and since the correlation between education is not 1, any influence for education is not likely to transmit 1:1 to income. Part of the luck that influences income may be genetic luck through unpredictable intermediate pathways. As

genetics influence near any trait for humans (Polderman et al. 2015), interests, hobbies, and knowledge-gaining that at first is have no profit-ambition may be influenced by genetics. With time, many let their interests influence their career path, turning their interests into their livelihood. In such a setting, we should not be surprised that both genetic and nonshared environments influence income to a larger degree than how many books in the home your parents had when you were a child. Currently, we lack ample theories to describe such pathways, and ample empirical studies to shed more light on the sort of theoretical mechanisms I've suggested.

Norway, and more and more countries around the world, are characterized by dual-earner households becoming the norm (although with frequent part-time work among young mothers). Subsidized, high-quality child care is available. The result is that in 2020 93 % of children in preschool ages 1-5 attain some kind of child care center (Statistics Norway 2021). Roughly half of the child's awake hours are spent without her/his parents, and some of those hours are spent commuting or getting ready to go to and from the child care center, hardly the time to foster cognitive skills. Should we expect family environments to matter in such a context? Perhaps the family environment has a pervasive influence on children beyond where the long arm of the equalizing welfare state can reach, but I don't think this should be our "default hypothesis". Furthermore, is the family environment caused by the parents in such a setting? In a supplementary analysis in paper 5, we show that genetic nurture may be a sibling interaction effect, as it's the sibling's PGS that predict the phenotype when we include genetic nurture components for mothers, fathers, and siblings. Papers 3-5 show that sibling differences according to birth order exist in a plethora of contexts, it exists in immigrant families, in multipartnered families, and in adults and young children in Norway, across many birth cohorts. Perhaps the peers influencing child development have shifted from parents to siblings, and friends and employees in the child care centers?

In sum, we can expect education and income to be under different influences. Maybe sociologists should stop attributing the causal precedence to the social background in general, and family environments in particular? If we open our imagination to include a wider range of explanations, we could discover new mechanisms. For education, the social background paradigm is still highly relevant, but so are genetic influences. For income, the social background paradigm has less support, and genetic and non-shared environmental influences seem more influential.

From research question 1 we learn that social inequality can be expected to be caused by many other factors than the family environment and that we should attempt to study both genetic-, shared-, and non-shared environmental influences. By doing so, we will gain a broader understanding of how social inequality in socioeconomic outcomes comes about.

In answering research question 2, I took a deeper drill into nonshared environmental influences by considering one of the few structural nonshared environmental differences we know, birth order differences. From a meta-analysis on educational attainment, Branigan et al (2013), we learn that the overall nonshared environmental influence is 24 %, whilst the heritability is 40 % and the shared environmental influence is 36 %. In other words, the nonshared environmental component amounts to two-thirds of the shared environmental influence. This is a substantial amount, and bear in mind that twin estimates are based on siblings that are born on the same day, and often attain the same classes, and share many other environmental influences. For regular siblings, the sibling correlation is around 0,25-0,40 for education (Wiborg and Hansen 2018), and from Branigan et al (2013) we can assume that a large chunk of this correlation is due to similar genetics. The nonsimilarity in siblings is striking.

In the investigation of birth order differences, the clearest take-home message from papers 3-5 is how pervasive they are. Regardless of parental origin of origin, half- or full sibling status, PGS, birth weight, and birth length, birth order differences exist in educational attainment/achievement. Being caused environmentally, within the family, it's one of the most clear-cut structural social inequality generating mechanisms in the social sciences. Both sociological accounts of social inequality and behavioral genetic accounts of inequality risks being confounded from the other, as we saw with sociological accounts of causal influences from family environments being confounded by genetics and PGSs capturing environmental influences. Birth order differences are not confounded and offer a window to learn about mechanisms causing social inequality also in other contexts than within the family.

In this paragraph, I should have been able to answer what causes birth order differences. The last word on what causes birth order differences is not yet written. Paper 5 shows that they're not caused by genetic differences or maternal in utero environmental factors. Neither are they influenced by parental resources to a large degree. We show that they are present already in 5th grade in paper 5, and we can thus expect them to arise early in childhood. In paper 3, we show that they are influenced by parental country of origin, showing that there is something environmental going on. What happens in those families with

parental origin from India, Sri Lanka, and Vietnam, I'm unable to fully account for. In paper 4, we raise the possibility that mothers may be more important for birth order differences than fathers, but we provide no decisive evidence. Having worked on birth order differences over three papers and many years, I'm beginning to think that there's more to sibling interactions than previously acknowledged. The confluence theory (Zajonc 1976, 1983; Zajonc and Markus 1975) and the resource dilution theory (Blake 1981, 1986, 1989; Gibbs, Workman, and Downey 2016) are the two competing theories on birth order differences. Of the two, the resource dilution theory often receives the most support, partly because the confluence theory has some very specific prediction on when birth order differences no longer matter (i.e. spacing of 11 ± 2 years), which are easy to test statistically and thus easy to dismiss on empirical grounds.

In sum, for research question 2, we learn that by including nonshared environments, we can expand our understanding of social inequality. In particular, birth order differences are pervasive and of substantial magnitude for education, and may well be as important as shared environments for siblings once we take genetics into account. Thus far, we lack the overall understanding of how important the various elements are, as many sociologists drill down into their preferred theoretical explanation and stay there, like an ostrich.

Why Genetics and Nonshared Environments?

Should we care about genetics and nonshared environments as sociologists, or should we leave these factors to others? Together, I hope the papers in this thesis show that considering heterogeneous effects of family background encompassing genetics and nonshared environmental influences opens up for a wider understanding in sociology for social inequality. In regards to nonshared environmental influences, few sociologists take an active stance for or against them. It is for the most part ignored as a knowledge gap that we do not know how to address and pose fewer problems to keep ignoring as it's something many disciplines share, and thus neighboring disciplines do not pose a "threat" to sociological findings. Genetics, on the other hand, is considered in many disciplines, and regularly raised as a potential confounder for sociological accounts of social inequality mechanisms. Here, I will briefly give some reasons why integrating genetics and nonshared environments are of importance for sociology as a discipline, with the main emphasis on genetics.

First, if we accept that humans are a mammalian species with Darwinian and Mendelian principles of evolution and genetics, there is no basis for denying that this may in

theory influence us at the population level, as well as individual level. By ignoring genetics, we are ignoring a part of reality, a part of the truth. In a discipline where many use individual-level outcomes to assess population-level phenomena, we would not benefit from ignoring an important part of individual-level make-up, their genetics, for the population-level outcome of interest. Quite simply, it gives a more realistic understanding of human behavior (Harden and Koellinger 2020).

Second, as long as sociological research into inequality concerns how it comes about, i.e. probing mechanisms generating social inequality, it would make no sense to ignore genetics or nonshared environmental influences. If sociologists are not willing to admit that genetics matter for sociological socio-economic outcomes, would they say the same on psychopathology and health-related outcomes? Would sociologists admit that autism, Down syndrome, cancer (i.e. BRCA1-2: BReast CAncer gene 1 and 2), spinal muscular atrophy, Huntington's disease, and other proximate human traits are influenced to a large degree by genetics, or in many cases determined by them? If so, and I'm assuming that most sociologists would, this will influence socio-economic attainments too, as many of these conditions and diseases make attaining socio-economic goods difficult. Note that diseases from rare mutations are highly likely to occur only in one sibling, and end up as nonshared "environment" in twin studies. Ignoring this part of the population would hamper our findings, as we are not including them. Here, I have left out the fact that genetically sensitive studies find a large association between genetic influences and socio-economic outcomes such as education and income, which does not overlap with the genetic influence for psychopathology. If we are interested in *social* influence, it would make sense to include genetic information if as nothing more than controls.

Third, combining genetics and environmental explanations opens up many possibilities. Currently, sociologists often analyze social variables and treat them as having homogenous in the intergenerational transmission of socio-economic outcomes. As an example, in contrast, behavioral genetics have concepts for three types of parental-child relations for child development in its introductory textbook, *genotype-environment correlations* (Knopik et al. 2017:117). For instance, they describe active genotype-environment correlation, which happens when individuals select, modify, construct or reconstruct experiences that correlate with their genetic propensities. To make this example a bit more tuned to the mainstream sociologist, replace "genetic propensities" with interests. Assume interests are partly innate to a child and let the example be musical ability. If a child

is musically talented and attends a school with musical instruments available, teachers may pick up the talent and encourage further development of it. If the talented kids' sibling attends the same school and the same musical instruments are available, but the innate talent is not at the same level, the sibling may not be encouraged to the same degree to pursue musical environments. In this vein, environments mediate genetic propensities.

The question then, would be if any given (parent) sociologist contests that many parents attempt to help children pursue their interests as far as they can afford economically and have time, more or less regardless of what they are, as long as they are not considered harmful? The same holds for preschools and schools. And if such active genotype-environment correlations are theoretically meaningful, would they not be of interest to sociologists? Childrearing practices have changed a lot over the past decades toward a more child-centric approach to development. As almost all human traits are partly heritable, parenting practices and child care institutions have may have increased the influence from genetics. Furthermore, in dual-earner households which Norwegian public policy actively encourages, children spend less time in the household than perhaps ever before, and childrearing is partly done in institutional settings like schools and preschools. Just by opening up the theoretical possibility that genetics and childrearing practices may interact in a myriad of ways, sociology can learn more about how social inequality differences arise (Breinholt and Conley 2020). Thus far, large swaths of sociology ignore such interdependencies, continuing instead on a black-box-like scenario with dated theories from different eras. Similarly, if siblings are not that similar in socio-economic attainments, wouldn't the causes of dissimilarity be interesting for sociology?

Fourth, public perception of sociology could decrease. Educated mothers with multiple children are the best guessers of how high heritability estimates are (Willoughby et al. 2018). If sociology continues to assert that genetics does not matter, it's in stark contrast to what other research finds, but also to what lay people rightfully believe to be true. If sociology continuous to ignore genetics, the general public can be left to think that sociology is something of an "ideological factory".

Fifth, following principles of vertical integration of science (Aakvaag 2010), sociology is high up on the integration ladder. Ignoring findings from lower levels and pretending to have a complete world view that is clearly in contrast to what fields in lower-ends of the integration-ladder finds could pose big problems for sociology. As others have argued, the findings from behavioral genetics are a challenge from a close-neighbor on the integration-

ladder, psychology, rather than biology (Freese 2008). Genes did not eat your homework, but something in your psychological predispositions may influence you to do so, and part of that influence may be genetic. Should sociologists attempt to simply ignore this, we risk losing credibility within academic circles too, following Stein Bråtens model monopolies (Bråten 2000) and Habermasian principles of the unforced force of the better argument. Instead, if sociology wants to be involved in the broad scientific community, sociology should glance to what neighboring fields find, read what seems of relevance to the discipline, and then criticize, incorporate, or explicitly ignore those findings after evaluating them. Then, sociology would maintain its credibility by the power of better arguments and its rightly earned status as a broad social science of many uses.

Sixth, by doing what I just described under point five, other disciplines become aware of sociology's strengths too. To be fair, a small smorgasbord of sociologists have been involved in sociogenomic research for a long time already (i.e. Conley and Fletcher 2017; Eckland 1967; Freese, Li, and Wade 2003; Jencks et al. 1972; Mills and Tropf 2020; Nielsen 1994; Scarr and Weinberg 1978; Udry 1995). They, and others in the social sciences, have provided invaluable insights into (behavioral) genetics. Sociologists' insights into genetics continue to this day. For example, the latest molecular genetic advance, polygenic scores, which I utilize in this thesis, have been shown to pick up environmental variation (Morris et al. 2020). It led a leading geneticist to state: "I didn't at first want to learn about sociology, but now I realize I have to. I have no choice" (Mills and Tropf 2020 footnote number 9). Sociologists have a long history of being a fragmented field with floating boundaries, which is regarded as one of its weaknesses as sociology may be colonized by other fields. However, it's also one of its strengths, as genetics can with relative ease be incorporated into the discipline, as it is already so diverse. Furthermore, sociologists master the structural aspects of societies with a long-standing history of dealing with structures and agents at all conceptual levels. This proves a benefit when geneticists without such training realize that their methods are not free from environmental confounding.

The Future is Now

Throughout this thesis, I have argued that incorporating genetics and paying closer attention to nonshared environments offer a broader perspective for research into social inequality. I don't think sociology can ignore genetics anymore. The sociogenomic era is here (Conley and Fletcher 2017; Freese 2018; Mills and Tropf 2020), and it's increasingly clear that genes have

real causal influences on many of the outcomes sociologists working with individual-level outcomes care about (Belsky et al. 2018; Freese 2008; Polderman et al. 2015; Turkheimer 2000). In the coming years, the accessibility of social science datasets with molecular genetic data will skyrocket due to lowered cost of collecting such data, and increased interest across many scientific disciplines, and the potential for use in health care. For now, sociogenomic studies typically use the educational attainment GWAS (Lee et al. 2018) to create a PGS for use to predict socio-economic outcomes. But there is rapid development in both methods (Eilertsen et al. 2021; Howe et al. 2021; Mills and Tropf 2020) and traits under consideration for GWAS, opening up for PGS' for many other outcomes of interest to social scientists. Income is one such outcome, where preliminary results show that there is overlap between SNPs that are associated with educational attainment and income, but the overlap is only 50 % (Hyeokmoon et al. 2020). Thus, if results hold and sample size and prediction increase similar to what happened between the iterations of the educational attainment GWAS' (Lee et al. 2018; Okbay et al. 2016; Rietveld et al. 2013), we will have a highly predictive PGS for income too. In the coming years, we will have many more PGSs to use for social science, many more datasets (with trio designs), and many other methodological advances and theoretical insights.

The lack of interest in sociology for genetics means we're late, but it's not too late to join the party. Twin studies have shown that shared environmental influences are low, and heritability high for most traits. Results from studies applying molecular genetic approaches are unable to replicate the same heritability estimates, they are far lower, known as the missing heritability problem (Maher 2008; Manolio et al. 2009). Why that is still unclear. It may be due to rare variants in the genome having large effects, and Whole-Genome Sequencing may solve some of the problems like it has done for height (Yang et al. 2015). However, it may also be that for complex traits like education and other traits sociologists care about twin studies have overestimated heritability due to assortative mating, gene-environment correlations, and reasons not fully understood. Once more, family data is hailed as a potential solution to try to unpack this puzzle (Young 2019), and agency-structure-like understandings are pivotal.

A Kodak moment is a cultural meme known as an occasion suitable for memorializing with a photograph. Kodak-moment as a cultural meme is sure to be lost on the younger generations, as Kodak lost its market position due to physical photographs no longer being profitable as digital photographs took over. Less known is the fact that Kodak invented the

digital photograph and patented many of the components used in digital cameras today. Kodak decided that digital photographs were a threat to their business model of printed photos. Sociology coined agency-structure, and we could have included genetics in our theories a long time ago. Genetics may seem threatening, but we will not know how much genetics is needed in sociology before we have trialed and errored.

References

- Aakvaag, Gunnar C. 2010. "Samfunnsvitenskapen Og Den Naturalistiske Utfordringen." *Nytt Norsk Tidsskrift* (3):264–78.
- Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25(3):153–72. doi: 10.1257/jep.25.3.153.
- Anon. 2017. "Population and Population Change Statistics - Statistics Explained." *Eurostat*. Retrieved November 16, 2017 (http://ec.europa.eu/eurostat/statistics-explained/index.php/Population_and_population_change_statistics).
- Baier, Tina, and Volker Lang. 2019. "The Social Stratification of Environmental and Genetic Influences on Education: New Evidence Using a Register-Based Twin Sample." *Sociological Science* 6:143–71. doi: 10.15195/v6.a6.
- Barclay, Kieron. 2014. "Birth Order and Intelligence : A Within- Family Analysis Using Swedish Military Conscription Data." *Stockholm Research Reports in Demography* 14:1–21.
- Barclay, Kieron. 2015. "Birth Order and Educational Attainment: Evidence from Fully Adopted Sibling Groups." *Intelligence* 48:109–22. doi: 10.1016/j.intell.2014.10.009.
- Barclay, Kieron, Martin Hällsten, and Mikko Myrskylä. 2016. *Birth Order and College Major in Sweden*. Rostock.
- Barclay, Kieron J., Torkild H. Lyngstad, and Dalton C. Conley. 2021. "The Production of Inequalities within Families and across Generations: The Intergenerational Effects of Birth Order on Educational Attainment." *European Sociological Review* 1–19. doi: 10.1093/esr/jcab005.
- Barrabés, Nuria, and Greta Kjølstad Østli. 2017. *Norwegian Standard Classification of Education 2016*. Oslo-Kongsvinger.
- Barth, Daniel, Nicholas W. Papageorge, and Kevin Thom. 2020. "Genetic Endowments and Wealth Inequality." *Journal of Political Economy* 128(4):1474–1522. doi: 10.1086/705415.
- Bates, Timothy C., Brion S. Maher, Sarah E. Medland, Kerrie McAloney, Margaret J. Wright, Narelle K. Hansell, Kenneth S. Kendler, Nicholas G. Martin, and Nathan A. Gillespie. 2018. "The Nature of Nurture: Using a Virtual-Parent Design to Test Parenting Effects on Children's Educational Attainment in Genotyped Families." *Twin Research and Human Genetics* 21(2):73–83. doi: 10.1017/thg.2018.11.
- Beebee, Helen, Christopher Hitchcock, and Peter Menzies. 2010. "Introduction." Pp. 1–16 in *The Oxford Handbook of Causation*.
- Behrman, J. R., and P. Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics* 4(3 Pt. 2):121–50. doi: 10.1086/298124.
- Behrman, J R, and P. Taubman. 1986. "Birth Order, Schooling, and Earnings." *Journal of Labor Economics* 4(3):121–50. doi: 10.1086/298124.
- Belsky, Daniel W., Benjamin W. Domingue, Robbee Wedow, Louise Arseneault, Jason D. Boardman, Avshalom Caspi, Dalton Conley, Jason M. Fletcher, Jeremy Freese, Pamela

- Herd, Terrie E. Moffitt, Richie Poulton, Kamil Sicinski, Jasmin Wertz, and Kathleen Mullan Harris. 2018. "Genetic Analysis of Social-Class Mobility in Five Longitudinal Studies." *Proceedings of the National Academy of Sciences* 115(31):E7275–84. doi: 10.1073/pnas.1801238115.
- Benjamin, Daniel J., David Cesarini, Christopher F. Chabris, Edward L. Glaeser, I. David, Vilmundur Gu, Tamara B. Harris, Lenore J. Launer, Shaun Purcell, Albert Vernon Smith, Magnus Johannesson, Patrik K. E. Magnusson, P. Jonathan, Nicholas A. Christakis, Craig S. Atwood, Benjamin Hebert, Jeremy Freese, Robert M. Hauser, Taissa S. Hauser, Alexander Grankvist, Christina M. Hultman, and Paul Lichtenstein. 2012. "The Promises and Pitfalls of Genoeconomics." *Annu Rev Econom* 4:627–62. doi: 10.1146/annurev-economics-080511-110939.The.
- Bertoni, Marco, and Giorgio Brunello. 2016. "Later-Borns Don't Give Up: The Temporary Effects of Birth Order on European Earnings." *Demography* 53:449–70. doi: 10.1007/s13524-016-0454-1.
- Black, Sandra E., Paul J. Devereux, Petter Lundborg, and Kaveh Majlesi. 2020. "Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviours." *Review of Economic Studies* 87(4):1683–1725. doi: 10.1093/restud/rdz038.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education*." *Quarterly Journal of Economics* 120(2):669–700. doi: 10.1162/0033553053970179.
- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2011. "Older and Wiser? Birth Order and IQ of Young Men." *CESifo Economic Studies* 57:103–20. doi: 10.1093/cesifo/ifq022.
- Blake, Judith. 1981. "Family Size and the Quality of Children." *Demography* 18(4):421–42.
- Blake, Judith. 1986. "Number of Siblings, Family Background, and the Process of Educational Attainment." *Social Biology* 33(1–2):5–21. doi: 10.1080/19485565.1986.9988618.
- Blake, Judith. 1989. *Family Size and Achievement*. Berkeley: University of California Press.
- Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. First Free. New York: The Free Press.
- Bonesrønning, Hans, and Sofia Sandgren Massih. 2011. "Birth Order Effects on Young Students' Academic Achievement." *Journal of Socio-Economics* 40:824–32. doi: 10.1016/j.socec.2011.08.010.
- Branigan, Amelia R., Kenneth J. Mccallum, and Jeremy Freese. 2013. "Variation in the Heritability of Educational Attainment: An International Meta-Analysis." *Social Forces* 92(1):109–40. doi: 10.1093/sf/sot076.
- Bråten, Stein. 2000. *Modellmakt Og Altersentriske Spedbarn : Essays on Dialogue in Infant and Adult*. Bergen: Sigma.
- Bratsberg, Bernt, and Ole Rogeberg. 2018. "Flynn Effect and Its Reversal Are Both Environmentally Caused." *Proceedings of the National Academy of Sciences of the United States of America* 115(26):6674–78. doi: 10.1073/pnas.1718793115.

- Breinholt, Asta, and Dalton Conley. 2020. "Child-Driven Parenting: Differential Early Childhood Investment by Offspring Genotype." *NBER Working Paper* (28217).
- Brenøe, Anne Ardila, and Ramona Molitor. 2017. "Birth Order and Health of Newborns: What Can We Learn from Danish Registry Data?" *Journal of Population Economics* 31(2):363–95. doi: 10.1007/s00148-017-0660-1.
- Burks, Barbara Stoddard. 1928. "THE RELATIVE INFLUENCE OF NATURE AND NURTURE UPON MENTAL DEVELOPMENT ; A COMPARATIVE RESEMBLANCE AND TRUE PARENT-TRUE CHILD RESEMBLANCE." *Yearbook of the National Society for the Study of Education* 27:219–316.
- Chabris, Christopher F., James J. Lee, David Cesarini, Daniel J. Benjamin, and David I. Laibson. 2015. "The Fourth Law of Behavior Genetics." *Current Directions in Psychological Science* 24(4):304–12. doi: 10.1177/0963721415580430.
- Christakis, Nicholas A. 2020. *Apollo's Arrow*. Little, Brown Spark.
- Conley, Dalton. 2005. *The Pecking Order*. New York: Vintage Books.
- Conley, Dalton, and Neil G. Bennett. 2000. "Is Biology Destiny ? Birth Weight and Life Chances." *American Sociological Review* 65(3):458–67.
- Conley, Dalton, Benjamin Domingue, David Cesarini, Christopher Dawes, Cornelius Rietveld, and Jason Boardman. 2015. "Is the Effect of Parental Education on Offspring Biased or Moderated by Genotype?" *Sociological Science* 2:82–105. doi: 10.15195/v2.a6.
- Conley, Dalton, and Jason Fletcher. 2017. *The Genome Factor*. Princeton: Princeton University Press.
- Conley, Dalton, Emily Rauscher, Christopher Dawes, Patrik K. E. Magnusson, and Mark Siegal L. 2013. "The Equal Environments Assumption in the Post-Genomic Age: Using Misclassified Twins to Estimate Bias in Heritability Models." *Behavior Genetics* 43(4). doi: 10.3386/w16711.
- Conley, Dalton, Emily Rauscher, Christopher Dawes, Patrik K. E. Magnusson, and Mark L. Siegal. 2013. "Heritability and the Equal Environments Assumption: Evidence from Multiple Samples of Misclassified Twins." *Behavior Genetics* 43(5):415–26. doi: 10.1007/s10519-013-9602-1.
- Conley, Dalton, Mark L. Siegal, Benjamin W. Domingue, Kathleen Mullan Harris, Matthew B. McQueen, and Jason D. Boardman. 2014. "Testing the Key Assumption of Heritability Estimates Based on Genome-Wide Genetic Relatedness." *Journal of Human Genetics* 59(6):342–45. doi: 10.1038/jhg.2014.14.
- Domingue, B. W., J. Fletcher, D. Conley, and J. D. Boardman. 2014. "Genetic and Educational Assortative Mating among US Adults." *Proceedings of the National Academy of Sciences* 111(22):7996–8000. doi: 10.1073/pnas.1321426111.
- Domingue, Benjamin W., Daniel W. Belsky, Jason M. Fletcher, Dalton Conley, Jason D. Boardman, and Kathleen Mullan Harris. 2018. "The Social Genome of Friends and Schoolmates in the National Longitudinal Study of Adolescent to Adult Health." *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1711803115.
- Duncan, Otis Dudley. 1968. "Ability and Achievement." *Biodemography and Social Biology*

- 15(1):1–11. doi: 10.1080/19485565.1968.9987747.
- Eckland, Bruce K. 1967. “Genetics and Sociology : A Reconsideration.” *American Sociological Review* 32(2):173–94.
- Eilertsen, Espen Moen, Eshim Shahid Jami, Tom A. McAdams, Laurie J. Hannigan, Alexandra S. Havdahl, Per Magnus, David M. Evans, and Eivind Ystrom. 2021. “Direct and Indirect Effects of Maternal, Paternal, and Offspring Genotypes: Trio-GCTA.” *Behavior Genetics* 51(2):154–61. doi: 10.1007/s10519-020-10036-6.
- England, Paula. 2010. “The Gender Revolution: Uneven and Stalled.” *Gender and Society* 24(2):149–66. doi: 10.1177/0891243210361475.
- Engzell, P., and F. C. Tropf. 2019. “Heritability of Education Rises with Intergenerational Mobility.” *Proceedings of the National Academy of Sciences of the United States of America* 116(51):25386–88. doi: 10.31235/osf.io/yeq68.
- Engzell, Per. 2019. “What Do Books in the Home Proxy For? A Cautionary Tale.” *Sociological Methods & Research*. doi: 10.1177/0049124119826143.
- Esping-Andersen, Gøsta. 1990. “The Three Worlds of Welfare Capitalism.”
- Esping-Andersen, Gøsta. 2015. “Welfare Regimes and Social Stratification.” *Journal of European Social Policy* 25(1):124–34. doi: 10.1177/0958928714556976.
- Fagereng, Andreas, Magne Mogstad, and Marte Rønning. 2021. “Why Do Wealthy Parents Have Wealthy Children?” *Journal of Political Economy* 129(3):000–000. doi: 10.1086/712446.
- Felson, Jacob. 2014. “What Can We Learn from Twin Studies? A Comprehensive Evaluation of the Equal Environments Assumption.” *Social Science Research* 43:184–99. doi: 10.1016/j.ssresearch.2013.10.004.
- Figlio, D., K. Karbownik, and K. G. Salvanes. 2016. *Education Research and Administrative Data*. Vol. 5. 1st ed. Elsevier B.V.
- Figlio, David N., Jeremy Freese, Krzysztof Karbownik, and Jeffrey Roth. 2017. “Socioeconomic Status and Genetic Influences on Cognitive Development.” *Proceedings of the National Academy of Sciences* 114(51):13441–46. doi: 10.1073/pnas.1708491114.
- Foster, John W., and Stanley J. Archer. 1979. “Birth Order and Intelligence: An Immunological Interpretation.” *Perceptual and Motor Skills* (48):79–93. doi: 10.2466/pms.1979.48.1.79.
- Freese, Jeremy. 2008. “Genetics and the Social Science Explanation of Individual Outcomes.” *American Journal of Sociology* 114(S1):S1–35. doi: 10.1086/592208.
- Freese, Jeremy. 2018. “The Arrival of Social Science Genomics.” *Contemporary Sociology: A Journal of Reviews* 47(5):524–36. doi: 10.1177/0094306118792214a.
- Freese, Jeremy, and Yu Han Jao. 2017. “Shared Environment Estimates for Educational Attainment: A Puzzle and Possible Solutions.” *Journal of Personality* 85(1):79–89. doi: 10.1111/jopy.12226.
- Freese, Jeremy, Jui-Chung Allen Li, and Lisa D. Wade. 2003. “The Potential Relevance of Biology to Social Inquiry.” *Annual Review of Sociology* 29(1):233–56. doi: 10.1146/annurev.soc.29.010202.100012.

- Freese, Jeremy, and Brian Powell. 2003. "Tilting at Windmills : Rethinking Sociological Responses to Behavioral Genetics." *Journal of Health and Social Behavior* 44(2):130–35.
- Freese, Jeremy, and Sara Shostak. 2009. "Genetics and Social Inquiry." *Annual Review of Sociology* 35(1):107–28. doi: 10.1146/annurev-soc-070308-120040.
- Garip, Filiz. 2020. "What Failure to Predict Life Outcomes Can Teach Us." *Proceedings of the National Academy of Sciences of the United States of America* 117(15):8234–35. doi: 10.1073/pnas.2003390117.
- Gibbs, Benjamin G., Joseph Workman, and Douglas B. Downey. 2016. "The (Conditional) Resource Dilution Model: State- and Community-Level Modifications." *Demography* 53(3):723–48. doi: 10.1007/s13524-016-0471-0.
- Giddens, Anthony, and Philip W. Sutton. 2017. *Sociology*. 8. ed. edited by P. W. Sutton. Cambridge: Polity.
- Grätz, Michael. 2018. "Competition in the Family: Inequality between Siblings and the Intergenerational Transmission of Educational Advantage." *Sociological Science* 5:246–69. doi: 10.15195/v5.a11.
- Grätz, Michael, and Martin Kolk. 2019. "Sibling Similarity in Income : A Life Course Perspective Sibling Similarity in Income : A Life Course Perspective." *Stockholm Research Reports in Demography 2014: 14*.
- Guhin, Jeffrey, Jessica McCrory Calarco, and Cynthia Miller-Idriss. 2021. "Whatever Happened to Socialization?" *Annual Review of Sociology*. doi: 10.1126/science.353.6305.1198.
- de Haan, Monique. 2010. "Birth Order, Family Size and Educational Attainment." *Economics of Education Review* 29:576–88. doi: 10.1016/j.econedurev.2009.10.012.
- De Haan, Monique, Erik Plug, and José Rosero. 2014. "Birth Order and Human Capital Development Evidence from Ecuador." *The Journal of Human Resources* 49(2):359–92.
- Haave, Per. 2000. *Sterilisering Av Tatere 1934-1977 - En Historisk Undersøkelse Av Lov Og Praksis*.
- Haider, Steven, and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4):1308–20. doi: 10.1257/aer.96.4.1308.
- Hamer, D. H., L. Sirota, Daniel Laurison, Dawn Dow, and Carolyn Chernoff. 2000. "Beware the Chopsticks Gene." *Molecular Psychiatry* 5(1):11–13. doi: 10.1038/sj.mp.4000662.
- Harden, K Paige. 2021. "'Reports of My Death Were Greatly Exaggerated': Behavior Genetics in the Postgenomic Era." *Annual Review of Psychology* 72(1):37–60. doi: 10.1146/annurev-psych-052220-103822.
- Harden, K. Paige. 2021. *The Genetic Lottery - Why DNA Matters for Social Equality*. Princeton University Press.
- Harden, K. Paige, Benjamin W. Domingue, Daniel W. Belsky, Jason D. Boardman, Robert Crosnoe, Margherita Malanchini, Michel Nivard, Elliot M. Tucker-Drob, and Kathleen Mullan Harris. 2020. "Genetic Associations with Mathematics Tracking and Persistence in Secondary School." *Npj Science of Learning* 5(1):1. doi: 10.1038/s41539-020-0060-2.

- Harden, K. Paige, and Philipp D. Koellinger. 2020. "Using Genetics for Social Science." *Nature Human Behaviour*. doi: 10.1038/s41562-020-0862-5.
- Härkönen, Juho. 2014. "Birth Order Effects on Educational Attainment and Educational Transitions in West Germany." *European Sociological Review* 30(2):166–79. doi: 10.1093/esr/jct027.
- Harris, Jennifer R., Per Magnus, and Kristian Tambs. 2006. "The Norwegian Institute of Public Health Twin Program of Research: An Update." *Twin Research and Human Genetics* 9(6):858–64. doi: 10.1375/183242706779462886.
- Hart, Sara A., Callie Little, and Elsje van Bergen. 2021. "Nurture Might Be Nature: Cautionary Tales and Proposed Solutions." *Npj Science of Learning* 6(1):1–12. doi: 10.1038/s41539-020-00079-z.
- Havari, Enkelejda, and Marco Savegnago. 2020. "The Intergenerational Effects of Birth Order on Education." *Journal of Population Economics*. doi: 10.1007/s00148-020-00810-5.
- Haworth, Simon, Ruth Mitchell, Laura Corbin, Kaitlin H. Wade, Tom Dudding, Ashley Budu-Aggrey, David Carslake, Gibran Hemani, Lavinia Paternoster, George Davey Smith, Neil Davies, Daniel J. Lawson, and Nicholas J. Timpson. 2019. "Apparent Latent Structure within the UK Biobank Sample Has Implications for Epidemiological Analysis." *Nature Communications* 10(1):1–9. doi: 10.1038/s41467-018-08219-1.
- Heath, Andrew C., K. Berg, L. J. Eaves, M. H. Solaas, L. A. Corey, Jon Martin Sundet, P. Magnus, and W. E. Nance. 1985. "Education Policy and the Heritability of Educational Attainment." *Nature* 314(25):734–36.
- Herd, Pamela, Jeremy Freese, Kamil Sicinski, Benjamin W. Domingue, Kathleen Mullan Harris, Caiping Wei, and Robert M. Hauser. 2019. "Genes, Gender Inequality, and Educational Attainment." *American Sociological Review* 84(6):1069–98. doi: 10.1177/0003122419886550.
- Hermansen, Are Skeie. 2016. "Moving Up or Falling Behind? Intergenerational Socioeconomic Transmission among Children of Immigrants in Norway." *European Sociological Review* 32(5):675–89. doi: 10.1093/esr/jcw024.
- Hill, W. David, Neil M. Davies, Stuart J. Ritchie, Nathan G. Skene, Julien Bryois, Steven Bell, Emanuele Di Angelantonio, David J. Roberts, Shen Xueyi, Gail Davies, David C. M. Liewald, David J. Porteous, Caroline Hayward, Adam S. Butterworth, Andrew M. McIntosh, Catharine R. Gale, and Ian J. Deary. 2019. "Genome-Wide Analysis Identifies Molecular Systems and 149 Genetic Loci Associated with Income." *Nature Communications* 10(1):5741. doi: 10.1038/s41467-019-13585-5.
- Hivert, Valentin, Julia Sidorenko, Florian Rohart, Michael E. Goddard, Jian Yang, Naomi R. Wray, Loic Yengo, and Peter M. Visscher. 2021. "Estimation of Non-Additive Genetic Variance in Human Complex Traits from a Large Sample of Unrelated Individuals." *The American Journal of Human Genetics* 1–13. doi: 10.1016/j.ajhg.2021.02.014.
- Horwitz, Allan V., Tami M. Videon, Mark F. Schmitz, and Diane Davis. 2003. "Rethinking Twins and Environments : Possible Social Sources for Assumed Genetic Influences in Twin Research." *Journal of Health and Social Behavior* 44(2):111–29.
- Howe, Laurence J., Geetha Chittoor, Penelope A. Lind, Michel G. Nivard, Tim T. Morris, Yunzhang Wang, Marianne Nygaard, Scott M. Ratliff, Chandra A. Reynolds, Jared V

- Balbona, F. Lawrence, Christina C. Dahm, Archie Campbell, Harry Campbell, Christopher R. Bauer, Scott D. Gordon, Zhengming Chen, Yongkang Kim, K. Paige Harden, Luke M. Evans, Alexandra Havdahl, David J. Porteous, W. David Hill, Margaret J. Wright, Nicholas G. Martin, Jennifer A. Smith, Caroline Hayward, Daniel J. Benjamin, and Ben Brumpton. 2021. “Within-Sibship GWAS Improve Estimates of Direct Genetic Effects.” *BioRxiv*. doi: 10.1101/2021.03.05.433935.
- Hugh-Jones, David, Karin J. H. Verweij, Beate St. Pourcain, and Abdel Abdellaoui. 2016. “Assortative Mating on Educational Attainment Leads to Genetic Spousal Resemblance for Polygenic Scores.” *Intelligence* 59:103–8. doi: 10.1016/j.intell.2016.08.005.
- Hyeokmoon, Kweon, Daphne Oluwaseun Martschenko, Kathryn Paige Harden, Thomas A. Diprete, and Philipp D. Koellinger. 2020. “Genetic Fortune : Winning or Losing Education , Income , and Health.”
- Hyttinen, Ari, Pekka Ilmakunnas, Edvard Johansson, and Otto Toivanen. 2019. “Heritability of Lifetime Earnings.” *The Journal of Economic Inequality*. doi: 10.1007/s10888-019-09413-x.
- Isungset, Martin Arstad. 2015. “Fødselsrekkefølgeeffekter i Utdanning: En Kvantitativ Studie Om Norskfødte Med Innvandrereforeldre.” University of Oslo.
- Isungset, Martin Arstad, Tina Baier, and Torkild Hovde Lyngstad. 2021. “Genetic Influences on Lifetime Income Increases with Gender Equality.” *SocArXiv*. doi: 10.31235/osf.io/e5wjk.
- Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- Jonsson, Hakon, Erna Magnusdottir, Hannes P. Eggertsson, Olafur A. Stefansson, Gudny A. Arnadottir, Ogmundur Eiriksson, Florian Zink, Einar A. Helgason, Ingileif Jonsdottir, Arnaldur Gylfason, Adalbjorg Jonasdottir, Aslaug Jonasdottir, Doruk Beyter, Thora Steingrimsdottir, Gudmundur L. Norddahl, Olafur Th Magnusson, Gisli Masson, Bjarni V Halldorsson, Unnur Thorsteinsdottir, Agnar Helgason, Patrick Sulem, Daniel F. Gudbjartsson, and Kari Stefansson. 2021. “Differences between Germline Genomes of Monozygotic Twins.” *Nature Genetics* 53(1):27–34. doi: 10.1038/s41588-020-00755-1.
- Kantarevic, Jasmin, and Stephane Mechoulam. 2006. “Birth Order, Educational Attainment, and Earnings: An Investigation Using the PSID.” *Journal of Human Resources* 755–77.
- Klepp, Ingun Grimstad. 2018. “Husmor [Housewives].” *Store Norske Leksikon*.
- Knopik, Valerie S., Jenae M. Neiderhiser, John C. DeFries, and Robert Plomin. 2017. *Behavioral Genetics*. 7th editio. New York: Worth Publishers.
- Kong, Augustine, Michael L. Frigge, Gudmar Thorleifsson, Hreinn Stefansson, Alexander I. Young, Florian Zink, Gudrun A. Jonsdottir, Aysu Okbay, Patrick Sulem, Gisli Masson, Daniel F. Gudbjartsson, Agnar Helgason, Gyda Bjornsdottir, Unnur Thorsteinsdottir, and Kari Stefansson. 2017. “Selection against Variants in the Genome Associated with Educational Attainment.” *Proceedings of the National Academy of Sciences* 114(5):E727–32. doi: 10.1073/pnas.1612113114.
- Kong, Augustine, Gudmar Thorleifsson, Michael L. Frigge, Bjarni Vilhjalmsen, Alexander I. Young, Thorgeir E. Thorgeirsson, Stefania Benonisdottir, Asmundur Oddsson, Bjarni V.

- Halldorsson, Gisli Masson, Daniel F. Gudbjartsson, Agnar Helgason, Gyda Bjornsdottir, Unnur Thorsteinsdottir, and Kari Stefansson. 2018. "The Nature of Nurture: Effects of Parental Genotypes." *Science* 359(January):424–28. doi: 10.1101/219261.
- Kravdal, Øystein. 2019. "Taking Birth Year into Account When Analysing Effects of Maternal Age on Child Health and Other Outcomes: The Value of a Multilevelmultiprocess Model Compared to a Sibling Model." *Demographic Research* 40(May):1249–90. doi: 10.4054/DemRes.2019.40.43.
- Kristensen, Petter, and Tor Bjerkedal. 2007. "Explaining the Relation Between Birth Order and Intelligence." *Science* 316(June):1717–1717. doi: 10.1126/science.1141493.
- Kristensen, Petter, and Tor Bjerkedal. 2010. "Educational Attainment of 25 Year Old Norwegians According to Birth Order and Gender." *Intelligence* 38:123–36. doi: 10.1016/j.intell.2009.08.003.
- Lander, Eric S., Lauren M. Linton, Bruce Birren, Chad Nusbaum, Michael C. Zody, Jennifer Baldwin, Keri Devon, Ken Dewar, Michael Doyle, William Fitzhugh, Roel Funke, Diane Gage, Katrina Harris, Andrew Heaford, John Howland, Lisa Kann, Jessica Lehoczy, Rosie Levine, Paul McEwan, Kevin McKernan, James Meldrim, Jill P. Mesirov, Cher Miranda, William Morris, Jerome Naylor, Christina Raymond, Mark Rosetti, Ralph Santos, Andrew Sheridan, Carrie Sougnez, Nicole Stange-Thomann, Nikola Stojanovic, Aravind Subramanian, Dudley Wyman, Jane Rogers, John Sulston, Rachael Ainscough, Stephan Beck, David Bentley, John Burton, Christopher Clee, Nigel Carter, Alan Coulson, Rebecca Deadman, Panos Deloukas, Andrew Dunham, Ian Dunham, Richard Durbin, Lisa French, Darren Grafham, Simon Gregory, Tim Hubbard, Sean Humphray, Adrienne Hunt, Matthew Jones, Christine Lloyd, Amanda McMurray, Lucy Matthews, Simon Mercer, Sarah Milne, James C. Mullikin, Andrew Mungall, Robert Plumb, Mark Ross, Ratna Shownkeen, Sarah Sims, Robert H. Waterston, Richard K. Wilson, Ladeana W. Hillier, John D. McPherson, Marco A. Marra, Elaine R. Mardis, Lucinda A. Fulton, Asif T. Chinwalla, Kymberlie H. Pepin, Warren R. Gish, Stephanie L. Chissoe, Michael C. Wendl, Kim D. Delehaunty, Tracie L. Miner, Andrew Delehaunty, Jason B. Kramer, Lisa L. Cook, Robert S. Fulton, Douglas L. Johnson, Patrick J. Minx, Sandra W. Clifton, Trevor Hawkins, Elbert Branscomb, Paul Predki, Paul Richardson, Sarah Wenning, Tom Slezak, Norman Doggett, Jan Fang Cheng, Anne Olsen, Susan Lucas, Christopher Elkin, Edward Uberbacher, Marvin Frazier, Richard A. Gibbs, Donna M. Muzny, Steven E. Scherer, John B. Bouck, Erica J. Sodergren, Kim C. Worley, Catherine M. Rives, James H. Gorrell, Michael L. Metzker, Susan L. Naylor, Raju S. Kucherlapati, David L. Nelson, George M. Weinstock, Yoshiyuki Sakaki, Asao Fujiyama, Masahira Hattori, Tetsushi Yada, Atsushi Toyoda, Takehiko Itoh, Chiharu Kawagoe, Hidemi Watanabe, Yasushi Totoki, Todd Taylor, Jean Weissenbach, Roland Heilig, William Saurin, Francois Artiguenave, Philippe Brottier, Thomas Bruls, Eric Pelletier, Catherine Robert, Patrick Wincker, André Rosenthal, Matthias Platzer, Gerald Nyakatura, Stefan Taudien, Andreas Rump, Douglas R. Smith, Lynn Doucette-Stamm, Marc Rubenfield, Keith Weinstock, Mei Lee Hong, Joann Dubois, Huanming Yang, Jun Yu, Jian Wang, Guyang Huang, Jun Gu, Leroy Hood, Lee Rowen, Anup Madan, Shizen Qin, Ronald W. Davis, Nancy A. Federspiel, A. Pia Abola, Michael J. Proctor, Bruce A. Roe, Feng Chen, Huaqin Pan, Juliane Ramser, Hans Lehrach, Richard Reinhardt, W. Richard McCombie, Melissa De La Bastide, Neilay Dedhia, Helmut Blöcker, Klaus Hornischer, Gabriele Nordsiek, Richa Agarwala, L. Aravind, Jeffrey A. Bailey, Alex Bateman, Serafim Batzoglou, Ewan Birney, Peer Bork, Daniel G. Brown, Christopher B. Burge, Lorenzo Cerutti, Hsiu Chuan Chen, Deanna Church, Michele Clamp, Richard R. Copley, Tobias

- Doerks, Sean R. Eddy, Evan E. Eichler, Terrence S. Furey, James Galagan, James G. R. Gilbert, Cyrus Harmon, Yoshihide Hayashizaki, David Haussler, Henning Hermjakob, Karsten Hokamp, Wonhee Jang, L. Steven Johnson, Thomas A. Jones, Simon Kasif, Arek Kasprzyk, Scot Kennedy, W. James Kent, Paul Kitts, Eugene V. Koonin, Ian Korf, David Kulp, Doron Lancet, Todd M. Lowe, Aoife McLysaght, Tarjei Mikkelsen, John V. Moran, Nicola Mulder, Victor J. Pollara, Chris P. Ponting, Greg Schuler, Jörg Schultz, Guy Slater, Arian F. A. Smit, Elia Stupka, Joseph Szustakowki, Danielle Thierry-Mieg, Jean Thierry-Mieg, Lukas Wagner, John Wallis, Raymond Wheeler, Alan Williams, Yuri I. Wolf, Kenneth H. Wolfe, Shiaw Pyng Yang, Ru Fang Yeh, Francis Collins, Mark S. Guyer, Jane Peterson, Adam Felsenfeld, Kris A. Wetterstrand, Richard M. Myers, Jeremy Schmutz, Mark Dickson, Jane Grimwood, David R. Cox, Maynard V. Olson, Rajinder Kaul, Christopher Raymond, Nobuyoshi Shimizu, Kazuhiko Kawasaki, Shinsei Minoshima, Glen A. Evans, Maria Athanasiou, Roger Schultz, Aristides Patrinos, Michael J. Morgan, Pieter de Jong, Joseph J. Catanese, Kazutoyo Osoegawa, Hiroaki Shizuya, Sangdun Choi, and Yu Jin Chen. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409:860–921. doi: 10.1038/35087627.
- Larsen, Erling Røed. 2001. "The Norwegian Economy 1900-2000 : From Rags to Riches." *Economic Survey* 4.
- Lee, James J., Robbee Wedow, Aysu Okbay, Edward Kong, Omeed Maghzian, Meghan Zacher, Tuan Anh Nguyen-Viet, Peter Bowers, Julia Sidorenko, Richard Karlsson Linnér, Mark Alan Fontana, Tushar Kundu, Chanwook Lee, Hui Li, Ruoxi Li, Rebecca Royer, Pascal N. Timshel, Raymond K. Walters, Emily A. Willoughby, Loïc Yengo, Maris Alver, Yanchun Bao, David W. Clark, Felix R. Day, Nicholas A. Furlotte, Peter K. Joshi, Kathryn E. Kemper, Aaron Kleinman, Claudia Langenberg, Reedik Mägi, Joey W. Trampush, Shefali Setia Verma, Yang Wu, Max Lam, Jing Hua Zhao, Zhili Zheng, Jason D. Boardman, Harry Campbell, Jeremy Freese, Kathleen Mullan Harris, Caroline Hayward, Pamela Herd, Meena Kumari, Todd Lencz, Jian'an Luan, Anil K. Malhotra, Andres Metspalu, Lili Milani, Ken K. Ong, John R. B. Perry, David J. Porteous, Marylyn D. Ritchie, Melissa C. Smart, Blair H. Smith, Joyce Y. Tung, Nicholas J. Wareham, James F. Wilson, Jonathan P. Beauchamp, Dalton C. Conley, Tõnu Esko, Steven F. Lehrer, Patrik K. E. Magnusson, Sven Oskarsson, Tune H. Pers, Matthew R. Robinson, Kevin Thom, Chelsea Watson, Christopher F. Chabris, Michelle N. Meyer, David I. Laibson, Jian Yang, Magnus Johannesson, Philipp D. Koellinger, Patrick Turley, Peter M. Visscher, Daniel J. Benjamin, and David Cesarini. 2018. "Gene Discovery and Polygenic Prediction from a Genome-Wide Association Study of Educational Attainment in 1.1 Million Individuals." *Nature Genetics* 50(8):1112–21. doi: 10.1038/s41588-018-0147-3.
- Lesthaeghe, Ron J. 2010. "The Unfolding Story of Transition." *Population and Development Review* 36(2):211–51. doi: 10.1111/j.1728-4457.2010.00328.x.
- Li, Jun Z., Devin M. Absher, Hua Tang, Audrey M. Southwick, Amanda M. Casto, Sohini Ramachandran, Howard M. Cann, Gregory S. Barsh, Marcus Feldman, Luigi L. Cavalli-sforza, and Richard M. Myers. 2008. "Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation." *Science* 319(February):1100–1105.
- Lichtenstein, Paul, Nancy L. Pedersen, and G. E. McClearn. 1992. "Origins of Individual Differences in Occupational Status and Educational Level." *Acta Sociologica* 35:13–31.
- Liu, Hexuan. 2018. "Social and Genetic Pathways in Multigenerational Transmission of Educational Attainment." *American Sociological Review* 83(2):278–304. doi:

10.1177/0003122418759651.

Lundberg, Ian. 2020. "Does Opportunity Skip Generations? Reassessing Evidence From Sibling and Cousin Correlations." *Demography* 57(4):1193–1213. doi: 10.1007/s13524-020-00880-w.

Lundberg, Ian, Rebecca Johnson, and M. Brandon Stewart. 2021. "What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory." *American Sociological Review*.

Magnus, Per, Charlotte Birke, Kristine Vejrup, Anita Haugan, Elin Alsaker, Anne Kjersti Daltveit, Marte Handal, Margaretha Haugen, Gudrun Høiseth, Gun Peggy Knudsen, Liv Paltiel, Patricia Schreuder, Kristian Tambs, Line Vold, and Camilla Stoltenberg. 2016. "Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa)." *International Journal of Epidemiology* 45(2):382–88. doi: 10.1093/ije/dyw029.

Magnus, Per, Lorentz M. Irgens, Kjell Haug, Wenche Nystad, Rolv Skjærven, Camilla Stoltenberg, Elin Alsaker, Leiv S. Bakketeig, Anne Kjersti Daltveit, Merete Eggesbø, Jorid Eide, Bodolf Hareide, Margaretha Haugen, Ragnhild Hovengen, Kari Kveim Lie, Rolv Terje Lie, Helle M. Meltzer, Roy Nilsen, Rannveig Nordhagen, Liv Paltiel, Ted Reichborn-Kjennerud, Kjersti Skjold Rønningen, Patricia Schreuder, Nina Stensrud, Arild Sunde, Grethe S. Tell, Stein Emil Vollset, and Johannes Wiik. 2006. "Cohort Profile: The Norwegian Mother and Child Cohort Study (MoBa)." *International Journal of Epidemiology* 35(5):1146–50. doi: 10.1093/ije/dyl170.

Maher, Brendan. 2008. "Personal Genomes: The Case of the Missing Heritability." *Nature* 456(7218):18–21. doi: 10.1038/456018a.

Manolio, Teri A., Francis S. Collins, Nancy J. Cox, David B. Goldstein, Lucia A. Hindorff, David J. Hunter, Mark I. McCarthy, Erin M. Ramos, Lon R. Cardon, Aravinda Chakravarti, Judy H. Cho, Alan E. Guttmacher, Augustine Kong, Leonid Kruglyak, Elaine Mardis, Charles N. Rotimi, Montgomery Slatkin, David Valle, Alice S. Whittemore, Michael Boehnke, Andrew G. Clark, Evan E. Eichler, Greg Gibson, Jonathan L. Haines, Trudy F. C. MacKay, Steven A. McCarroll, and Peter M. Visscher. 2009. "Finding the Missing Heritability of Complex Diseases." *Nature* 461(7265):747–53. doi: 10.1038/nature08494.

Mazumder, Bhashkar. 2011. "Family and Community Influences on Health and Socioeconomic Status : Sibling Correlations Over the Life Course." *The B.E. Journal of Economic Analysis & Policy* 11(3).

Mills, Melinda C., Nicola Barban, and Felix C. Tropf. 2020. *An Introduction to Statistical Genetic Data Analysis*. edited by N. Barban and F. C. Tropf. Cambridge, Massachusetts: The MIT Press.

Mills, Melinda C., and Felix C. Tropf. 2020. "Sociology, Genetics and the Coming of Age of Sociogenomics." *Annual Review of Sociology* 1–29. doi: 10.1146/annurev-soc-121919-054756.

Monden, Christiaan, Gilles Pison, and Jeroen Smits. 2021. "Twin Peaks: More Twinning in Humans than Ever Before." *Human Reproduction* 1–8. doi: 10.1093/humrep/deab029.

Mönkediek, Bastian, Wiebke Schulz, Harald Eichhorn, and Martin Diwald. 2020. "Is There Something Special about Twin Families? A Comparison of Parenting Styles in Twin and Non-Twin Families." *Social Science Research* 90(July). doi:

10.1016/j.ssresearch.2020.102441.

- Morris, Tim T., Neil M. Davies, Gibran Hemani, and George Davey Smith. 2020. "Population Phenomena Inflate Genetic Associations of Complex Social Traits." *Science Advances* 6(16). doi: 10.1126/sciadv.aay0328.
- Mukherjee, Siddhartha. 2015. *The Gene: An Intimate History*. London: Penguin Group.
- Muslimova, Dilnoza, Hans Van Kippersluis, Cornelius A. Rietveld, Stephanie Von Hinke, and S. Fleur W. Meddens. 2020. "Dynamic Complementarity in Skill Production : Evidence from Genetic Endowments and Birth Order." *Tinbergen Institute Discussion Paper* 1–46. doi: 10.2139/ssrn.3748468.
- Nielsen, François. 1994. "Sociobiology and Sociology." *Annual Review of Sociology* 20:267–303.
- Nielsen, François. 2006. "Achievement and Ascription in Educational Attainment: Genetic and Environmental Influences on Adolescent Schooling." *Social Forces* 85(1):193–216. doi: 10.1353/sof.2006.0135.
- Nielsen, François. 2018. "Genes and Status Achievement." Pp. 1–27 in *Oxford Handbook of Evolution, Biology and Society*, edited by R. L. Hopcroft. Oxford University Press.
- Nielsen, François, and J. Micah Roos. 2015. "Genetics of Educational Attainment and the Persistence of Privilege at the Turn of the 21st Century." *Social Forces* 94(2):535–61. doi: 10.1093/sf/sov080.
- Nielsen, Torben Hviid, Arve Monsen, and Tore Tennøe. 2000. *Livets Tre Og Kodenes Kode : Fra Genetikk Til Bioteknologi : Norge 1900-2000*. edited by A. Monsen, T. Tennøe, and A. Eriksen. Oslo: Gyldendal akademisk.
- Nilsen, Thomas S., Ingunn Brandt, Per Magnus, and Jennifer R. Harris. 2012. "The Norwegian Twin Registry." *Twin Research and Human Genetics* 15(6):775–80. doi: 10.1017/thg.2012.57.
- Nilsen, Thomas S., Gun Peggy Knudsen, Kristina Gervin, Ingunn Brandt, Espen Roysamb, Kristian Tambs, Ragnhild Orstavik, Robert Lyle, Ted Reichborn-Kjennerud, Per Magnus, and Jennifer R. Harris. 2013. "The Norwegian Twin Registry from a Public Health Perspective: A Research Update." *Twin Research and Human Genetics* 16(1):285–95. doi: 10.1017/thg.2012.117.
- Okbay, Aysu, Jonathan P. Beauchamp, Mark Alan Fontana, James J. Lee, Tune H. Pers, Cornelius A. Rietveld, Patrick Turley, Guo Bo Chen, Valur Emilsson, S. Fleur W. Meddens, Sven Oskarsson, Joseph K. Pickrell, Kevin Thom, Pascal Timshel, Ronald De Vlaming, Abdel Abdellaoui, Tarunveer S. Ahluwalia, Jonas Bacelis, Clemens Baumbach, Gyda Bjornsdottir, Johannes H. Brandsma, Maria Pina Concas, Jaime Derringer, Nicholas A. Furlotte, Tessel E. Galesloot, Giorgia Grotto, Richa Gupta, Leanne M. Hall, Sarah E. Harris, Edith Hofer, Momoko Horikoshi, Jennifer E. Huffman, Kadri Kaasik, Ioanna P. Kalafati, Robert Karlsson, Augustine Kong, Jari Lahti, Sven J. Van Der Lee, Christiaan Deleeuw, Penelope A. Lind, Karl Oskar Lindgren, Tian Liu, Massimo Mangino, Jonathan Marten, Evelin Mihailov, Michael B. Miller, Peter J. Van Der Most, Christopher Oldmeadow, Antony Payton, Natalia Pervjakova, Wouter J. Peyrot, Yong Qian, Olli Raitakari, Rico Rueedi, Erika Salvi, Borge Schmidt, Katharina E. Schraut, Jianxin Shi, Albert V. Smith, Raymond A. Poot, Beate St Pourcain, Alexander Teumer, Gudmar Thorleifsson, Niek Verweij, Dragana Vuckovic, Juergen

Wellmann, Harm Jan Westra, Jingyun Yang, Wei Zhao, Zhihong Zhu, Behrooz Z. Alizadeh, Najaf Amin, Andrew Bakshi, Sebastian E. Baumeister, Ginevra Biino, Klaus Bønnelykke, Patricia A. Boyle, Harry Campbell, Francesco P. Cappuccio, Gail Davies, Jan Emmanuel De Neve, Panos Deloukas, Ilja Demuth, Jun Ding, Peter Eibich, Lewin Eisele, Niina Eklund, David M. Evans, Jessica D. Faul, Mary F. Feitosa, Andreas J. Forstner, Ilaria Gandin, Bjarni Gunnarsson, Bjarni V. Halldórsson, Tamara B. Harris, Elizabeth G. Holliday, Andrew C. Heath, Lynne J. Hocking, Georg Homuth, Michael A. Horan, Jouke Jan Hottenga, Philip L. De Jager, Peter K. Joshi, Astanand Jugessur, Marika A. Kaakinen, Mika Kähönen, Stavroula Kanoni, Liisa Keltigangas-Järvinen, Lambertus A. L. M. Kiemeny, Ivana Kolcic, Seppo Koskinen, Aldi T. Kraja, Martin Kroh, Zoltan Kutalik, Antti Latvala, Lenore J. Launer, Maël P. Lebreton, Douglas F. Levinson, Paul Lichtenstein, Peter Lichtner, David C. M. Liewald, Anu Loukola, Pamela A. Madden, Reedik Mägi, Tomi Mäki-Opas, Riccardo E. Marioni, Pedro Marques-Vidal, Gerardus A. Meddens, George McMahon, Christa Meisinger, Thomas Meitinger, Yusplitri Milaneschi, Lili Milani, Grant W. Montgomery, Ronny Myhre, Christopher P. Nelson, Dale R. Nyholt, William E. R. Ollier, Aarno Palotie, Lavinia Paternoster, Nancy L. Pedersen, Katja E. Petrovic, David J. Porteous, Katri Räikkönen, Susan M. Ring, Antonietta Robino, Olga Rostapshova, Igor Rudan, Aldo Rustichini, Veikko Salomaa, Alan R. Sanders, Antti Pekka Sarin, Helena Schmidt, Rodney J. Scott, Blair H. Smith, Jennifer A. Smith, Jan A. Staessen, Elisabeth Steinhagen-Thiessen, Konstantin Strauch, Antonio Terracciano, Martin D. Tobin, Sheila Ulivi, Simona Vaccargiu, Lydia Quaye, Frank J. A. Van Rooij, Cristina Venturini, Anna A. E. Vinkhuyzen, Uwe Völker, Henry Völzke, Judith M. Vonk, Diego Vozzi, Johannes Waage, Erin B. Ware, Gonneke Willemsen, John R. Attia, David A. Bennett, Klaus Berger, Lars Bertram, Hans Bisgaard, Dorret I. Boomsma, Ingrid B. Borecki, Ute Bültmann, Christopher F. Chabris, Francesco Cucca, Daniele Cusi, Ian J. Deary, George V. Dedoussis, Cornelia M. Van Duijn, Johan G. Eriksson, Barbara Franke, Lude Franke, Paolo Gasparini, Pablo V. Gejman, Christian Gieger, Hans Jörgen Grabe, Jacob Gratten, Patrick J. F. Groenen, Vilmundur Gudnason, Pim Van Der Harst, Caroline Hayward, David A. Hinds, Wolfgang Hoffmann, Elina Hyppönen, William G. Iacono, Bo Jacobsson, Marjo Riitta Järvelin, Karl Heinz Jöckel, Jaakko Kaprio, Sharon L. R. Kardia, Terho Lehtimäki, Steven F. Lehrer, Patrik K. E. Magnusson, Nicholas G. Martin, Matt Mcgue, Andres Metspalu, Neil Pendleton, Brenda W. J. H. Penninx, Markus Perola, Nicola Pirastu, Mario Pirastu, Ozren Polasek, Danielle Posthuma, Christine Power, Michael A. Province, Nilesh J. Samani, David Schlessinger, Reinhold Schmidt, Thorkild I. A. Sørensen, Tim D. Spector, Kari Stefansson, Unnur Thorsteinsdottir, A. Roy Thurik, Nicholas J. Timpson, Henning Tiemeier, Joyce Y. Tung, André G. Uitterlinden, Veronique Vitart, Peter Vollenweider, David R. Weir, James F. Wilson, Alan F. Wright, Dalton C. Conley, Robert F. Krueger, George Davey Smith, Albert Hofman, David I. Laibson, Sarah E. Medland, Michelle N. Meyer, Jian Yang, Magnus Johannesson, Peter M. Visscher, Toñu Esko, Philipp D. Koellinger, David Cesarini, and Daniel J. Benjamin. 2016. “Genome-Wide Association Study Identifies 74 Loci Associated with Educational Attainment.” *Nature* 533(7604):539–42. doi: 10.1038/nature17671.

Ørstavik, Ragnhild E., Nikolai Czajkowski, Espen Roysamb, Gun Peggy Knudsen, Kristian Tambs, and Ted Reichborn-Kjennerud. 2014. “Sex Differences in Genetic and Environmental Influences on Educational Attainment and Income.” *Twin Research and Human Genetics* 17(6):516–25. doi: 10.1017/thg.2014.68.

OWiD. 2021. *Real Disposable Household Income Indexed*.

Papageorge, Nicholas W., and Kevin Thom. 2019. “Genes, Education, and Labor Market

- Outcomes: Evidence from the Health and Retirement Study.” *Journal of the European Economic Association* 1–49. doi: 10.1093/jeea/jvz072.
- Parsons, Talcott. 1953. *The Social System*. London: Routledge & Kegan Paul.
- Parsons, Talcott. 1968. *The Structure of Social Action: A Study in Social Theory with Special Reference to a Group of Recent European Writers*. New York: The Free Press.
- Paul, L. A. 2010. “Counterfactual Theories.” Pp. 1–27 in *The Oxford Handbook of Causation*. Vol. 1, edited by H. Beebe, C. Hitchcock, and P. Menzies. Oxford University Press.
- Pearl, Judea, and Dana Mackenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. Great Britain: Allen Lane.
- Pekkarinen, Tuomas, Kjell G. Salvanes, and Matti Sarvimäki. 2017. “The Evolution of Social Mobility: Norway during the Twentieth Century.” *Scandinavian Journal of Economics* 119(1):5–33. doi: 10.1111/sjoe.12205.
- Pinker, Steven. 2002. *The Blank Slate - The Modern Denial of Human Nature*. New York: Penguin Books.
- Plomin, Robert, and Denise Daniels. 1987. “Why Are Children in the Same Family so Different from One Another?” *Behavioral and Brain Sciences* 10(1):1–16. doi: 10.1017/S0140525X00055941.
- Polderman, Tinca J. C., Beben Benyamin, Christiaan A. De Leeuw, Patrick F. Sullivan, Arjen Van Bochoven, Peter M. Visscher, and Danielle Posthuma. 2015. “Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies.” *Nature Genetics* 47(7):702–9. doi: 10.1038/ng.3285.
- Psillos, Stathis. 2010. *Regularity Theories*. Vol. 1. edited by H. Beebe, C. Hitchcock, and P. Menzies. Oxford University Press.
- Richardson, Ken, and Sarah Norgate. 2005. “The Equal Environments Assumption of Classical Twin Studies May Not Hold.” *British Journal of Educational Psychology* 75(3):339–50. doi: 10.1348/000709904X24690.
- Rietveld, C. A., S. E. Medland, J. Derringer, J. Yang, T. Esko, N. W. Martin, H. J. Westra, K. Shakhbazov, A. Abdellaoui, A. Agrawal, E. Albrecht, B. Z. Alizadeh, N. Amin, J. Barnard, S. E. Baumeister, K. S. Benke, L. F. Bielak, J. A. Boatman, P. A. Boyle, G. Davies, C. de Leeuw, N. Eklund, D. S. Evans, R. Ferhmann, K. Fischer, C. Gieger, H. K. Gjessing, S. Hagg, J. R. Harris, C. Hayward, C. Holzapfel, C. A. Ibrahim-Verbaas, E. Ingelsson, B. Jacobsson, P. K. Joshi, A. Jugessur, M. Kaakinen, S. Kanoni, J. Karjalainen, I. Kolcic, K. Kristiansson, Z. Kutalik, J. Lahti, S. H. Lee, P. Lin, P. A. Lind, Y. Liu, K. Lohman, M. Loitfelder, G. McMahon, P. M. Vidal, O. Meirelles, L. Milani, R. Myhre, M. L. Nuotio, C. J. Oldmeadow, K. E. Petrovic, W. J. Peyrot, O. Polasek, L. Quaye, E. Reinmaa, J. P. Rice, T. S. Rizzi, H. Schmidt, R. Schmidt, A. V. Smith, J. A. Smith, T. Tanaka, A. Terracciano, M. J. H. M. van der Loos, V. Vitart, H. Volzke, J. Wellmann, L. Yu, W. Zhao, J. Allik, J. R. Attia, S. Bandinelli, F. Bastardot, J. Beauchamp, D. A. Bennett, K. Berger, L. J. Bierut, D. I. Boomsma, U. Bultmann, H. Campbell, C. F. Chabris, L. Cherkas, M. K. Chung, F. Cucca, M. de Andrade, P. L. De Jager, J. E. De Neve, I. J. Deary, G. V. Dedoussis, P. Deloukas, M. Dimitriou, G. Eiriksdottir, M. F. Elderson, J. G. Eriksson, D. M. Evans, J. D. Faul, L. Ferrucci, M. E. Garcia, H. Gronberg, V. Gudnason, P. Hall, J. M. Harris, T. B. Harris, N. D. Hastie, A. C. Heath, D. G. Hernandez, W. Hoffmann, A. Hofman, R. Holle, E. G. Holliday, J. J.

- Hottenga, W. G. Iacono, T. Illig, M. R. Jarvelin, M. Kahonen, J. Kaprio, R. M. Kirkpatrick, M. Kowgier, A. Latvala, L. J. Launer, D. A. Lawlor, T. Lehtimäki, J. Li, P. Lichtenstein, P. Lichtner, D. C. Liewald, P. A. Madden, P. K. E. Magnusson, T. E. Makinen, Marco Masala, M. McGue, Andres Metspalu, A. Mielck, M. B. Miller, G. W. Montgomery, Sutapa Mukherjee, D. R. Nyholt, B. A. Oostra, L. J. Palmer, A. Palotie, B. W. J. H. Penninx, M. Perola, P. A. Peyser, M. Preisig, K. Raikonen, O. T. Raitakari, A. Realo, S. M. Ring, S. Ripatti, F. Rivadeneira, I. Rudan, A. Rustichini, V. Salomaa, A. P. Sarin, D. Schlessinger, R. J. Scott, H. Snieder, B. St Pourcain, J. M. Starr, J. H. Sul, I. Surakka, R. Svento, A. Teumer, H. Tiemeier, F. J. A. van Rooij, D. R. Van Wagoner, E. Vartiainen, J. Viikari, P. Vollenweider, J. M. Vonk, G. Waeber, D. R. Weir, H. E. Wichmann, E. Widen, G. Willemsen, J. F. Wilson, A. F. Wright, D. Conley, G. Davey-Smith, L. Franke, P. J. F. Groenen, A. Hofman, M. Johannesson, S. L. R. Kardia, R. F. Krueger, D. Laibson, N. G. Martin, M. N. Meyer, D. Posthuma, A. R. Thurik, N. J. Timpson, A. G. Uitterlinden, C. M. van Duijn, P. M. Visscher, D. J. Benjamin, D. Cesarini, and P. D. Koellinger. 2013. "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment." *Science* 340(6139):1467–71. doi: 10.1126/science.1235488.
- Rimfeld, Kaili, Eva Krapohl, Maciej Trzaskowski, Jonathan R. I. Coleman, Saskia Selzam, Philip S. Dale, Tonu Esko, Andres Metspalu, and Robert Plomin. 2018. "Genetic Influence on Social Outcomes during and after the Soviet Era in Estonia." *Nature Human Behaviour* 2(4):269–75. doi: 10.1038/s41562-018-0332-5.
- Rodgers, Joseph Lee. 2001. "What Causes Birth Order-Intelligence Patterns? The Admixture Hypothesis, Revived." *American Psychologist* 56(6/7):505–10. doi: 10.1037/0003-066X.56.6-7.505.
- Røed, Knut, and Oddbjørn Raaum. 2003. "Administrative Registers - Unexplored Reservoirs of Scientific Knowledge?" *Economic Journal* 113(488):258–81. doi: 10.1111/1468-0297.00134.
- Rønningen, Kjersti S., Liv Paltiel, Helle M. Meltzer, Rannveig Nordhagen, Kari K. Lie, Ragnhild Hovengen, Margaretha Haugen, Wenche Nystad, Per Magnus, and A. Jane. 2006. "The Biobank of the Norwegian Mother and Child Cohort Study: A Resource for the next 100 Years." *European Journal of Epidemiology* 21(8):619–25.
- Salganik, Matthew J., Ian Lundberg, Alexander T. Kindel, Caitlin E. Ahearn, Khaled Al-Ghoneim, Abdullah Almaatouq, Drew M. Altschul, Jennie E. Brand, Nicole Bohme Carnegie, Ryan James Compton, Debanjan Datta, Thomas Davidson, Anna Filippova, Connor Gilroy, Brian J. Goode, Eaman Jahani, Ridhi Kashyap, Antje Kirchner, Stephen McKay, Allison C. Morgan, Alex Pentland, Kivan Polimis, Louis Raes, Daniel E. Rigobon, Claudia V. Roberts, Diana M. Stanescu, Yoshihiko Suhara, Adaner Usmani, Erik H. Wang, Muna Adem, Abdulla Alhajri, Bedoor AlShebli, Redwane Amin, Ryan B. Amos, Lisa P. Argyle, Livia Baer-Bositis, Moritz Büchi, Bo-Ryehn Chung, William Eggert, Gregory Faletto, Zhilin Fan, Jeremy Freese, Tejomay Gadgil, Josh Gagné, Yue Gao, Andrew Halpern-Manners, Sonia P. Hashim, Sonia Hausen, Guanhua He, Kimberly Higuera, Bernie Hogan, Ilana M. Horwitz, Lisa M. Hummel, Naman Jain, Kun Jin, David Jurgens, Patrick Kaminski, Areg Karapetyan, E. H. Kim, Ben Leizman, Naijia Liu, Malte Möser, Andrew E. Mack, Mayank Mahajan, Noah Mandell, Helge Marahrens, Diana Mercado-Garcia, Viola Mocz, Katariina Mueller-Gastell, Ahmed Musse, Qiankun Niu, William Nowak, Hamidreza Omidvar, Andrew Or, Karen Ouyang, Katy M. Pinto, Ethan Porter, Kristin E. Porter, Crystal Qian, Tamkinat Rauf, Anahit

- Sargsyan, Thomas Schaffner, Landon Schnabel, Bryan Schonfeld, Ben Sender, Jonathan D. Tang, Emma Tsurkov, Austin van Loon, Onur Varol, Xiafei Wang, Zhi Wang, Julia Wang, Flora Wang, Samantha Weissman, Kirstie Whitaker, Maria K. Wolters, Wei Lee Woon, James Wu, Catherine Wu, Kengran Yang, Jingwen Yin, Bingyu Zhao, Chenyun Zhu, Jeanne Brooks-Gunn, Barbara E. Engelhardt, Moritz Hardt, Dean Knox, Karen Levy, Arvind Narayanan, Brandon M. Stewart, Duncan J. Watts, and Sara McLanahan. 2020. "Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration." *Proceedings of the National Academy of Sciences* 117(15):8398–8403. doi: 10.1073/pnas.1915006117.
- Sauder, Michael. 2020. "A Sociology of Luck." *Sociological Theory* 1–24. doi: 10.1177/0735275120941178.
- Scarr, Sandra, and Louise Carter-Saltzman. 1979. "Twin Method : Defense of a Critical Assumption." *Behavior Genetics* 9(6).
- Scarr, Sandra, and Richard A. Weinberg. 1978. "The Influence of " Family Background " on Intellectual Attainment." *American Sociological Review* 43(5):674–92.
- Schuster, Stephan C., Webb Miller, Aakrosh Ratan, Lynn P. Tomsho, Belinda Giardine, Lindsay R. Kasson, Robert S. Harris, Desiree C. Petersen, Fangqing Zhao, Ji Qi, Can Alkan, Jeffrey M. Kidd, Yazhou Sun, Daniela I. Drautz, Pascal Bouffard, Donna M. Muzny, Jeffrey G. Reid, Lynne V. Nazareth, Qingyu Wang, Richard Burhans, Cathy Riemer, Nicola E. Wittekindt, Priya Moorjani, Elizabeth A. Tindall, Charles G. Danko, Wee Siang Teo, Anne M. Buboltz, Zhenhai Zhang, Qianyi Ma, Arno Oosthuysen, Abraham W. Steenkamp, Hermann Oostuisen, Philippus Venter, John Gajewski, Yu Zhang, B. Franklin Pugh, Kateryna D. Makova, Anton Nekrutenko, Elaine R. Mardis, Nick Patterson, Tom H. Pringle, Francesca Chiaromonte, James C. Mullikin, Evan E. Eichler, Ross C. Hardison, Richard A. Gibbs, Timothy T. Harkins, and Vanessa M. Hayes. 2010. "Complete Khoisan and Bantu Genomes from Southern Africa." *Nature* 463(7283):943–47. doi: 10.1038/nature08795.
- Sesardic, Neven. 2005. *Making Sense of Heritability*.
- Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. "The Educational and Early Occupational Attainment Process." *American Sociological Review* 34(1):82–92.
- Silventoinen, Karri, Aline Jelenkovic, Reijo Sund, Antti Latvala, Chika Honda, Fujio Inui, Rie Tomizawa, Mikio Watanabe, Norio Sakai, Esther Rebato, Andreas Busjahn, Jessica Tyler, John L. Hopper, Juan R. Ordoñana, Juan F. Sánchez-Romera, Lucia Colodro-Conde, Lucas Calais-Ferreira, Vinicius C. Oliveira, Paulo H. Ferreira, Emanuela Medda, Lorenza Nisticò, Virgilia Toccaceli, Catherine A. Derom, Robert F. Vlietinck, Ruth J. F. Loos, Sisira H. Siribaddana, Matthew Hotopf, Athula Sumathipala, Fruhling Rijdsdijk, Glen E. Duncan, Dedra Buchwald, Per Tynelius, Finn Rasmussen, Qihua Tan, Dongfeng Zhang, Zengchang Pang, Patrik K. E. Magnusson, Nancy L. Pedersen, Anna K. Dahl Aslan, Amie E. Hwang, Thomas M. Mack, Robert F. Krueger, Matt McGue, Shandell Pahlen, Ingunn Brandt, Thomas S. Nilsen, Jennifer R. Harris, Nicholas G. Martin, Sarah E. Medland, Grant W. Montgomery, Gonneke Willemsen, Meike Bartels, Catharina E. M. van Beijsterveldt, Carol E. Franz, William S. Kremen, Michael J. Lyons, Judy L. Silberg, Hermine H. Maes, Christian Kandler, Tracy L. Nelson, Keith E. Whitfield, Robin P. Corley, Brooke M. Huibregtse, Margaret Gatz, David A. Butler, Adam D. Tarnoki, David L. Tarnoki, Hang A. Park, Jooyeon Lee, Soo Ji Lee, Joohon Sung, Yoshie Yokoyama, Thorkild I. A. Sørensen, Dorret I. Boomsma, and Jaakko Kaprio.

2020. "Genetic and Environmental Variation in Educational Attainment: An Individual-Based Analysis of 28 Twin Cohorts." *Scientific Reports* 10(1):1–11. doi: 10.1038/s41598-020-69526-6.
- Statistics Norway. 2021. "Barnehager [Kindergardens]." Retrieved April 7, 2021 (<https://www.ssb.no/barnehager>).
- Steelman, Lala Carr, Brian Powell, Regina Werum, and Scott Carter. 2002. "Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges." *Annual Review of Sociology* 28:243–69. doi: 10.1146/annurev.soc.28.111301.093304.
- Sundet, Jon Martin. 2014. "The Flynn Effect in Families: Studies of Register Data on Norwegian Military Conscripts and Their Families." *Journal of Intelligence* 2:106–18. doi: 10.3390/jintelligence20301.
- Swift, Adam. 2004. "Would Perfect Mobility Be Perfect?" *European Sociological Review* 20(1):1–11. doi: 10.1093/esr/20.1.1.
- Tambs, Kristian, Torbjørn Rønning, C. Prescott, Kenneth S. Kendler, Ted Reichborn-, Sverre Torgersen, and Jennifer R. Harris. 2010. "The Norwegian Institute of Public Health Twin Study of Mental Health: Examining Recruitment and Attrition Bias." *Twin Research and Human Genetics* 12(2):158–68. doi: 10.1375/twin.12.2.158.The.
- Tenikue, Michel, and Bertrand Verheyden. 2010. "Birth Order and Schooling: Theory and Evidence from Twelve Sub-Saharan Countries." *Journal of African Economies* 19(4):459–95. doi: 10.1093/jae/ejq013.
- Thuesen, Nils Petter, Geir Thorsnæs, and Sissel Røvik. 2019. "Norge [Norway]." *Store Norske Leksikon*.
- Turkheimer, Eric. 2000. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science* 9(5):160–64. doi: 10.1111/1467-8721.00084.
- Turkheimer, Eric. 2016. "Weak Genetic Explanation 20 Years Later." *Perspectives on Psychological Science* 11(1):24–28. doi: 10.1177/1745691615617442.
- Udry, J. Richard. 1995. "Sociology and Biology : What Biology Do Sociologists Need to Know?" *Social Forces* 73(4):1267–78.
- UNDP. 2018. "Human Development Indices and Indicators. 2018 Statistical Update." *United Nations Development Programme* 27(4):123.
- Visscher, Peter M., William G. Hill, and Naomi R. Wray. 2008. "Heritability in the Genomics Era - Concepts and Misconceptions." *Nature Reviews Genetics* 9(4):255–66. doi: 10.1038/nrg2322.
- Visscher, Peter M., Sarah E. Medland, Manuel A. R. Ferreira, Katherine I. Morley, Gu Zhu, Belinda K. Cornes, Grant W. Montgomery, and Nicholas G. Martin. 2006. "Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings." *PLoS Genetics* 2(3):0316–25. doi: 10.1371/journal.pgen.0020041.
- Vladasel, Theodor, Matthew J. Lindquist, Joeri Sol, and Mirjam van Praag. 2020. "On the Origins of Entrepreneurship: Evidence from Sibling Correlations." *Journal of Business Venturing* (March):106017. doi: 10.1016/j.jbusvent.2020.106017.
- Wagner, Sander, Diederik Boertien, and Mette Gørtz. 2020. "The Wealth of Parents: Trends

- Over Time in Assortative Mating Based on Parental Wealth.” *Demography*. doi: 10.1007/s13524-020-00906-3.
- Wiborg, Øyvind N., and Marianne N. Hansen. 2018. “The Scandinavian Model During Increasing Inequality: Recent Trends in Educational Attainment, Earnings and Wealth among Norwegian Siblings.” *Research in Social Stratification and Mobility* 56(November 2016):53–63. doi: 10.1016/j.rssm.2018.06.006.
- Willoughby, Emily A., Alan C. Love, Matt McGue, William G. Iacono, Jack Quigley, and James J. Lee. 2018. “Free Will, Determinism, and Intuitive Judgments About the Heritability of Behavior.” *Behavior Genetics* 49(2):136–53. doi: 10.1007/s10519-018-9931-1.
- Woodward, James F. 2010. “Agency and Interventionist Theories.” *The Oxford Handbook of Causation* (November):1–28. doi: 10.1093/oxfordhb/9780199279739.003.0012.
- Yang, Jian, Andrew Bakshi, Zhihong Zhu, Gibran Hemani, Anna A. E. Vinkhuyzen, Sang Hong Lee, Matthew R. Robinson, John R. B. Perry, Ilja M. Nolte, Jana V van Vliet-Ostaptchouk, Harold Snieder, Tonu Esko, Lili Milani, Reedik Mägi, Andres Metspalu, Anders Hamsten, Patrik K. E. Magnusson, Nancy L. Pedersen, Erik Ingelsson, Nicole Soranzo, Matthew C. Keller, Naomi R. Wray, Michael E. Goddard, and Peter M. Visscher. 2015. “Genetic Variance Estimation with Imputed Variants Finds Negligible Missing Heritability for Human Height and Body Mass Index.” *Nature Genetics* 47(10):1114–20. doi: 10.1038/ng.3390.
- Young, Alexander I. 2019. “Solving the Missing Heritability Problem” edited by J. Flint. *PLOS Genetics* 15(6):e1008222. doi: 10.1371/journal.pgen.1008222.
- Zaidi, Batool, and S. Philip Morgan. 2017. “The Second Demographic Transition Theory: A Review and Appraisal.” *Annual Review of Sociology* 43(1):473–92. doi: 10.1146/annurev-soc-060116-053442.
- Zajonc, R. B. 1976. “Family Configuration and Intelligence.” *Science* 192(4236):227–36. doi: 10.1126/science.192.4236.227.
- Zajonc, R. B. 1983. “Validating the Confluence Model.” *Psychological Bulletin* 93(3):457–80. doi: 10.1037//0033-2909.93.3.457.
- Zajonc, R. B., and Frank J. Sulloway. 2007. “The Confluence Model: Birth Order as a Within-Family or Between-Family Dynamic?” *Personality and Social Psychology Bulletin* 33(9):1187–94. doi: 10.1177/0146167207303017.
- Zajonc, R., and G. Markus. 1975. “Birth Order and Intellectual Development.” *Psychological Review* 82(1):74–88. doi: 10.1037/h0076229.

Paper 4:

Lillehagen, Mats, and Martin Arstad Isungset. 2020. “New Partner, New Order? Multipartnered Fertility and Birth Order Effects on Educational Achievement.”* Published in *Demography* 57(5):1625–46

*Supplementary materials not included here, please see online at *Demography*

New Partner, New Order? Multipartnered Fertility and Birth Order Effects on Educational Achievement

Mats Lillehagen¹ · Martin Arstad Isungset¹

Published online: 15 September 2020
© The Author(s) 2020

Abstract

A substantial amount of research shows that younger siblings perform worse than their older sisters and brothers in several socioeconomic outcomes, including educational achievement. Most of these studies examined stable families and excluded half-siblings. However, the increasing prevalence of multipartnered fertility implies that many children grow up in nonnuclear families. We examine whether there is evidence for birth order effects in this context, which offers an opportunity to test and potentially expand the explanatory scope of the two main theories on birth order effects. We use comprehensive Norwegian registry data to study siblings in the 1985–1998 cohorts born to mothers or fathers who parented children with at least two partners. We provide evidence for negative effects of birth order on lower secondary school grades in both cases. Children born to fathers displaying multipartnered fertility tend to have lower grades than older full siblings but perform more similarly or better compared with older half-siblings. For siblings born to mothers with the multipartnered fertility pattern, later-born siblings do worse in school compared with all older siblings. This indicates that negative birth order effects tend to operate either within or across sets of full siblings, depending on the sex of the parent displaying multipartnered fertility. We argue that these findings can be explained by a combination of resource dilution/confluence theory and sex differences in residential arrangements following union dissolutions. We also suggest an alternative interpretation: maternal resources could be more important for generating negative birth order effects.

Keywords Birth order · Siblings · Education · Multipartner fertility

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13524-020-00905-4>) contains supplementary material, which is available to authorized users.

✉ Mats Lillehagen
mats.lillehagen@sosgeo.uio.no

¹ Department of Sociology and Human Geography, University of Oslo, PO Box 1096 Blindern, 0317 Oslo, Norway

Introduction

Being later-born confers a negative effect on several outcomes, including educational performance and choices (Barclay et al. 2017; Black et al. 2005; De Haan 2010; Kristensen and Bjerkedal 2010), intelligence scores (Black et al. 2011; Boomsma et al. 2008; Kristensen and Bjerkedal 2007; Sulloway 2007), income and choice of profession (Björklund et al. 2004; Grinberg 2015; Kantarevic and Mechoulam 2006), as well as health and mortality (Barclay and Myrskylä 2014; Black et al. 2016).

The extant literature, although both methodically rigorous and empirically rich, generally limits its scope to studying stable nuclear families. This approach has clearly served as a good starting point for birth order research given that traditional family patterns are still the most prevalent in Western countries, including Norway (Lappegård and Rønsen 2013; Thomson 2014). However, an increasing proportion of children and adults find themselves in families characterized by greater levels of instability and complexity (Andersson 2004; Andersson and Kolk 2015; Cancian et al. 2011; Kalmijn 2015a). Half-siblings, stepsiblings, stepparents, and a larger number of less clear-cut family relations have become more common (Guzzo 2014; Lichter et al. 2010; Thomson 2014).

An important part of this picture is the increasing prevalence of multipartnered fertility—that is, parenting children with more than one partner during one’s reproductive history (Carlson and Furstenberg 2006; Lappegård and Rønsen 2013; Thomson et al. 2014). For example, estimates for representative samples or populations from the United States, Australia, Sweden, and Norway found that between 10% and 20% of mothers and fathers display this fertility behavior based on data of births occurring from the 1980s to the early 2000s (Thomson et al. 2014). In this study, we extend the existing literature on birth order effects by considering such complex fertility histories. We utilize comprehensive Norwegian registry data and study the association between birth order and educational achievement. Our main samples are restricted to siblings born to either fathers or mothers who have parented at least three children in total with more than one co-parent. We analyze two samples that include siblings who share either a father or a mother displaying multipartnered fertility.

We examine three research questions. First, we ask whether there are negative birth order effects on educational outcomes, measured by average grades at the end of compulsory education, in the context of complex fertility patterns. The answer is not obvious: older siblings are more likely to be affected by adverse consequences of union dissolution and other related factors (Bronte-Tinkew et al. 2009; Fomby and Osborne 2017; Sigle-Rushton et al. 2014; Tach et al. 2010) that could attenuate, cancel out, or even reverse negative birth order effects.

Second, we investigate whether birth order effects are associated with the birth order of the child within the same *set*—that is, among full siblings who share both parents, with the birth order of the child relative to all full *and* half-siblings, or both. This distinction, which becomes relevant when considering multipartnered fertility, allows us to answer questions such as whether being the overall firstborn to a parent displaying this fertility pattern is better in terms of school achievement compared with being the firstborn within a set of full siblings. More generally, we are able to ask whether birth order effects operate mainly *within* or *across* sets of full siblings.

Third, we ask whether these associations differ depending on the sex of the parent displaying multipartnered fertility. We expect such differences because residential arrangements and the frequency of parental contact tend to vary with sex for parents who display multipartnered fertility.

In a broader societal perspective, this study provides a better understanding of the intrafamily distribution of socioeconomic advantages and disadvantages associated with an increasingly frequent type of fertility behavior. Norway is a good starting point for such analyses given that it is often considered to be among the forerunners of the second demographic transition (Lappegård et al. 2011; Lesthaeghe 2010).

Theory and Previous Research

We begin by presenting the two dominant theories of birth order. Both theories emphasize the importance of children living together or at least interacting frequently with both parents, competing for resources, tutoring each other, and so forth (Black et al. 2005; Hertwig et al. 2002).¹

Resource Dilution Theory

The main idea behind the resource dilution theory is that the firstborn child has the exclusive attention and resources of the parents, whereas such resources are diluted for younger siblings who must share them with their older siblings (Downey 2001). An underlying assumption is that parents attempt to distribute their finite resources equally between the children. Given that early investments are more important (e.g., Carneiro et al. 2015), this leads to negative effects of birth order on socioeconomic outcomes. The theory has received much attention, and studies have tested it using detailed data on actual time spent with children (Pavan 2016; Price 2008), and quality time more specifically (Monfardini and See 2016; Price 2008). The importance of smoking during pregnancy (Lundberg et al. 2010), use of child health care services (Butikofer et al. 2015), and breastfeeding (Buckles and Kolka 2014) have also been investigated. Although the theory is generally well supported empirically, the questions of which type of resources matter for birth order effects and their relative importance remain contested (e.g., Monfardini and See 2016).

Confluence Theory

Another important and closely related theory is the confluence theory (Conley 2004; Zajonc 1983; Zajonc and Markus 1975). The central explanatory concept in this theory is referred to as the *average intellectual climate* of the family. With each additional child born into the family, the intellectual climate drops not only for the new child but also for the earlier-born. The intellectual family environment is therefore better for the

¹ Birth order effects have been linked to *in utero* immunological responses and other biological processes (Dewey and Cohen 2007; Foster and Archer 1979). However, studies attempting to distinguish *in utero* influences from postnatal environments have found that the former are not important (Barclay 2015; Brenøe and Molitor 2015; Kristensen and Bjerkedal 2007).

firstborn than the second-born, for instance: the firstborn benefitted from initially having the intellectual environment exclusively with his/her parents. The confluence theory also postulates a positive effect for older children tutoring their younger siblings (e.g., Silles 2010), which yields negative birth order effects. Although the empirical expectations associated with this theory are generally similar to those of resource-based theories, some of the predictions made are more specific. The spacing between births should play an important role in generating differences in outcomes by birth order, and it has been claimed that the negative effect of being born later will reverse at the age of 11 (Zajonc et al. 1979). Thus far, no decisive evidence exists, and some of these more specific predictions have not been supported in recent studies (Barclay and Kolk 2018; Bonesrønning and Massih 2011; Härkönen 2014; Lehmann et al. 2018).

The Two Theories and Multipartnered Fertility

The two theories suggest that regular interaction with both parents in a common family environment is the most important precondition for the existence of negative birth order effects. In most previous studies, excluding children with half-siblings made such interactions between siblings and both parents very likely to occur in the samples studied. Many (if not most) of the children in our samples, however, experienced union dissolutions. In some cases, they never lived with both parents simultaneously. Still, we argue that it is reasonable to expect negative birth order effects based on the two theories, but three key issues need to be accounted for to derive more specific predictions: (1) the distinction between overall and full biological birth order, (2) the significance of union dissolutions and family complexity, and (3) sex-differentiated parental behavior (discussed in the upcoming section on different expectations for the two samples). We will now discuss these in turn beginning with the first, which is crucial for evaluating the significance of the latter two.

Overall and Full Biological Birth Order

In studies of nuclear families, birth order is assigned based on the year of birth for all siblings who share the same mother *and* father. When considering birth order in the context of multipartnered fertility, however, there are multiple fathers and/or mothers involved. We distinguish between two types of birth order to provide an adequate descriptive overview and because the implications derived from the two theories discussed will differ depending on the type of birth order.

First, we define the *overall* birth order based on the year of birth for each child relative to all full and half-siblings with whom one shares the same father or mother—that is, all children from both *sets*. Second, we refer to the *full biological* birth order of each child based on a child's year of birth relative to other full siblings parented by the same mother *and* father—siblings *within* the same set. The latter definition of birth order is similar to those used in previous studies, whereas the first is novel.

An example will make this important distinction more transparent. Consider our first sample, which includes siblings born to mothers who have parented children with more than one partner. Say that a mother has four biological children: two with the first father and two with the second. These four children are assigned an overall birth order from 1 to 4 based on their year of birth in ascending order. For the full biological birth order,

however, the two oldest children in this example are assigned values of 1 and 2 based on their year of birth relative to each other. The last two born are also given the values of 1 and 2 because they are (respectively) first- and second-born in the mother's last set. In this example, the values for the two types of birth order are identical for the two oldest children (overall = 1, 2; full biological = 1, 2) but different for the two youngest (overall = 3, 4; full biological = 1, 2).

Union Dissolution and Family Complexity

In a nuclear family setting, parental fixed effects make it possible to provide a robust *ceteris paribus* estimate of birth order effects. In the context of multipartnered fertility, having a higher birth order is correlated with a lower probability of having experienced a divorce as well as a higher probability of having an older half-sibling. Further, children in the second set are more likely to experience greater family stability, which could lead to confounding of negative effects of *both* full biological and overall birth order on educational achievement.

One potential complication is related to the fact that children of the first set in both our samples have experienced a union dissolution at some point, typically before the first child/children of the second set is/are born. Union dissolutions affect children's educational achievements negatively (e.g., Steele et al. 2009). The occurrence of union dissolutions could affect birth order effects within each set (i.e., the full biological birth order effects) if this event affects children differently depending on their age (Sigle-Rushton et al. 2014). Previous studies taking both birth order and union dissolutions into account simultaneously for siblings who share both parents have found that this is generally not the case (Kantarevic and Mechoulam 2006; Kristensen and Bjerkedal 2010; Sigle-Rushton et al. 2014). Therefore, we do not expect union dissolutions to affect the magnitude of any full biological birth order effects on educational attainment.

Still, union dissolution and family complexity or instability may cause differences *between* sets of siblings, which could affect our estimate of the *overall* birth order effects. Union dissolutions are likely to affect children in the first set negatively given that these children will have experienced this event in almost all cases. Additionally, many researchers writing about multipartnered fertility and complex families have been concerned about the more general disadvantages associated with growing up in these types of situations, including worse educational outcomes (Bernardi and Radl 2014; Bronte-Tinkew et al. 2009; Fomby and Osborne 2017; Fomby et al. 2016; McLanahan et al. 2013; Tach et al. 2010). A common argument is that the complexity of such arrangements leads to considerable stress for both children and parents (Petren 2017). We argue that these negative effects are likely to work in the same direction as union dissolutions, affecting children of the first set negatively, because we expect the second union to provide a more stable environment (Beaujouan 2016; Poortman and Lyngstad 2007).

These factors therefore imply a positive correlation between school achievement and overall birth order and thus an association opposite of the typically negative birth order effect. This implies that the estimated magnitude of the overall birth order effect in this context will be biased downward (become less negative), and what we obtain is therefore an estimate of this effect net of these offsetting mechanisms.

Different Expectations for the Two Samples in Light of the Resource Dilution and Confluence Theories

In both the resource dilution and the confluence theory, parent-child interaction causes birth order effects. In our samples, union dissolutions have occurred, and the parent-child interaction is not the same as in nuclear families. Mothers and fathers will typically create a family environment together with their child(ren) as a couple before they split up and the child(ren) start to live with their mother, while the father may or may not still be involved. We expect this to have different implications for birth order effects in our two samples.

For children to *fathers* who display multipartnered fertility, siblings in each set typically grow up in parental environments that are separate most of the time. Although all children will typically spend some time with both parents, the children of the first set (in particular) will typically spend much of their upbringing with their mother and their full siblings (and occasionally stepfathers). From the point of view of the two theories, this implies a discontinuity compared with the nuclear family context. Initially, both parents are present, which resembles the nuclear family mechanism. In the father-based sample, we expect that this will mainly give rise to birth order effects between full siblings (i.e., full biological birth order effects). Although the father is present and contributes (both before and after the union dissolution) to some degree, his presence is not consistent in the same way as that of the children's mother, with whom they typically reside. In other words, because a father displaying multipartnered fertility spends less time distributing his resources between *all* his children simultaneously (because his pattern of investment varies across sets with time), we are less likely to see evidence for overall birth order effects in the father-based sample. However, the combined efforts of the father and the mother within each set will give rise to full biological birth order effects.

We have different expectations for the *mother* sample, based on the same premises. In this situation, both full and half-siblings will spend more time interacting in a common environment. Although it clearly varies across time whether both the father and mother (or stepparents) are present, the mother's continued presence provides a consistency that differs from that seen in the father-based sample. Importantly, this implies that the mechanisms specified in the two theories can operate for all full and half-siblings that share the same mother. In other words, although both mothers and fathers contribute in both samples, the residential situations allow for continuity across sets in this case that was not present in the other sample, therefore giving rise to *overall* birth order effects. We therefore expect birth order mechanisms to operate mainly *within* sets of full siblings in the father-based sample and *across* in the mother-based sample.

This argument is based on some important premises, which are supported in existing literature. Previous evidence from Norway for the period covered by our data (children in the 1985–1998 birth cohorts) indicates that mothers tend to live with all or most of their biological children in the overwhelming majority of cases in which the parents have separated or never lived together (Jensen and Clausen 1997; Lappegård and Rønsen 2013). In the far-less common case of children registered as living with their fathers, some type of joint physical custody is present more often than not, implying that the mother is typically consistently involved with all her children in this situation as

well (Kitterød et al. 2016). Furthermore, in Norway during the period studied, full paternal physical custody was almost non-existent, and joint physical custody was extremely rare when children were registered as living with their mothers. Jensen and Clausen (1997) documented a prevalence of approximately 4% of cases at the end of our study period. This was also the general finding in studies based on data from many other Western countries in this period (e.g., Baxter 2012; Cancian and Meyer 1998; Singer 2008; Sodermans et al. 2013; Tach et al. 2010). Also, although fathers are often still involved without physical custody, both Norwegian and international evidence indicate that parenting children with new partners is associated with less paternal involvement with older children born to a previous co-parent (Berger et al. 2012; Kalmijn 2015b; Kitterød and Lyngstad 2012; Manning and Smock 1999, Skjørten et al. 2007).

Our expectations are summarized in Table 1.

Data and Sample Restriction

We use comprehensive registry data that cover the complete population of Norway. We link information on individuals from several registries through a nonidentifiable and unique ID number and study the relationship between birth order and grades (*Grunnskolepoeng*, hereafter referred to as *school points*) at the end of compulsory education (lower secondary school). The population registry enables us to link all individuals to their parents and thereby to their full siblings and half-siblings. Information on school points was obtained from a registry of educational information, which contains data on nearly all pupils during the period under observation.

Our samples include all children of mothers (the mother-based sample) or fathers (the father-based sample) who parented at least three children in total with at least two partners. One or more of the co-parent(s) in each sample could also display multipartnered fertility. To estimate the full biological birth order, the same person must be the co-parent of at least two of these children. Information on lower secondary school points are measured at age 16 and are available from 2001 to 2014. This leaves us with information on the cohorts from 1985 to 1998. We therefore include children of fathers or mothers whose first child was born in or after 1985 and who have at least two siblings, which implies that we have information on school points for all older full and half-siblings for everyone in our sample.

Individuals with younger siblings who are not old enough to have finished lower secondary school in 2014 are included in the main samples. We exclude all siblings of children with missing information on school points. Fifth-born and later-born children

Table 1 Empirical expectations for each sample and birth order types based on the theoretical discussion

Sample	Birth Order Type	
	Full Biological	Overall
Father	Yes	Weak/unclear
Mother	Yes	Yes

were also excluded because of the very small number of observations in these categories. Neither exclusion affected our main findings.

Our main analyses use a within-father/within-mother design, which creates additional data demands. Because the fixed-effects estimations require variation in birth order between children of the same father, the same mother, and combinations of fathers and mothers, we exclude children of plural births. These restrictions leave us with 20,776 and 19,508 observations in the father-based and mother-based samples, respectively. Additionally, we include results from a nuclear family sample that uses the same criteria but excludes those with half-siblings, and which is limited to parents with four or fewer children ($N = 396,864$).² This sample therefore includes only those children of two parents who exclusively parented children with each other.

Statistical Approach

Explanatory Variables: Birth Order

This study uses a distinction in two types of birth order. First, we define the overall birth order based on the year of birth for each individual relative to all full and half-siblings with the same father (father sample) or mother (mother sample). We assign a value between 1 and 4 to each child. Overall birth order is included in our models as a set of four dummy variables for those who are firstborn (reference group), second-born, third-born, and fourth-born.

The second type of birth order is full biological birth order of each child, similar to typical measures of birth order used in within-family studies. We assign a value to the child based on his or her year of birth relative to other full siblings parented by the same mother *and* father. The measure is also included as a set of dummy variables with firstborn as the reference group, and we also include dummy variables for second-born and third-born.

Outcome Variable

School Points at the End of Compulsory Education

We measure educational success as the school point average at the completion of lower secondary school, which occurs at age 16. Lower secondary school comprises grades 8–10 and is mandatory in Norway. The school point variable combines the most important grades achieved throughout this level of schooling. Grades are awarded on a scale of 1 to 6, and both overall achievement grades and exam grades are included. The averages of all grades are calculated and then multiplied by 10 to obtain the value of the variable on a metric scale. The grading system was altered slightly in 2007, and a tendency for the average to rise somewhat during the period has been documented (Bakken and Elstad 2012). The inclusion of control variables for birth cohort enables us to adjust for such effects.

² Whenever we refer to differences or similarities between this sample and our main samples, such statements are based on pooling the samples and doing interaction tests (results are available upon request).

Covariates

In addition to our measures of birth order, we include a set of individual-level covariates, which have been shown to be important for the proper estimation of birth order effects and are therefore considered standard (see, e.g., Black et al. 2005; De Haan 2010; Härkönen 2014; Kristensen and Bjerkedal 2010).

Parents' Age at Birth

Subtracting the parent's year of birth from each child's year of birth gives us the parent's age at birth. We include this information as two sets of dummy variables with six categories (16–19 years old, 20–24 years old (the reference category), 25–29 years old, 30–34 years old, 35–39 years old, and 40 years or older).

Child's Year of Birth

Child's year of birth is included as a set of dummy variables for each year from 1985 to 1998, with the first year as the reference category. This variable serves as a standardizer of the grade scale as well as a cohort control.

Child's Gender

Because gender is known to be important to educational outcomes, the child's gender is included as a dummy variable to standardize the coefficients (0 = male, 1 = female).

Variables Included in Additional Analyses

We include measures for spacing (full samples) and cohabitation (available for children born from 1987 to 1998) in several additional analyses in the online appendix. We provide more details on these variables in the relevant tables.

Analytical Strategy

We estimate three main models with the grade average at the end of compulsory education as the outcome variable in all three. We use linear regression models estimated separately for each sample. Our approach enables us to apply father or mother fixed effects (depending on the sample), and we include control variables for the child's gender, year of birth, and the age at birth of both parents. We examine the relationship between school points and the overall birth order (Model 1) and the full biological birth order (Model 2). Model 3 is similar to the first two but includes *both* types of birth order, which allows us to examine the relative importance of each type given that the other is adjusted for. This strategy is possible because the two types of birth order have different values for some children in our sample, as indicated in the examples discussed earlier in connection with the definition of the birth order measures. We use a similar specification when we estimate birth order effects for the nuclear family sample, but the clustering in the fixed effects in this model are based on combinations of mothers' and fathers' IDs.

The Educational Attainment of the Co-parent

In most rigorous studies of birth order, effects are estimated based on a within-family design—that is, using fixed effects for combinations of mothers and fathers. This estimation strategy, however, is not applicable in all cases given our main samples and analytical approach because every mother or father is required to have parented children with at least two co-parents. It is possible that these co-parents differ systematically in characteristics that affect the educational success of their children through social or genetic transmission. This could potentially cause or confound birth order effects associated with the *overall* birth order in cases where half siblings are compared. To consider this issue, we first include a measure of the educational attainment of each co-parent as a covariate. This makes almost no difference to our estimates (see Tables A6 and A7 in the online appendix), and because the inclusion of this mediating variable could introduce endogeneity issues, we opt to not include it in our preferred models (Morgan and Winship 2015).

Results

Table 2 provides descriptive information about the most important variables included in our analyses. The descriptive results are quite similar for the two samples. However, fathers displaying multipartnered fertility parent slightly more children overall with more partners, as reflected in the slightly higher proportions in the upper birth order categories.

Because of the way our main samples were constructed, they include equal numbers of first-, second-, and third-born siblings when the *overall* birth order is considered. We also include information on the number of children per partner for the most common combinations to provide a better overview of our sample: 1-2 indicates one child with the first partner and two with the second, 2-1 indicates two children with the first partner and one with the second, and so on. These combinations are fairly equally distributed between the samples, although the 1-3 combination (one child with the first partner and three with the second) is slightly more common among fathers.

Figure 1 shows the association between school points and birth order for each combination of the number of children per partner. Tables A1 and A2 of the online appendix provide the full numerical results for all birth order coefficients.³ Although the estimates are somewhat statistically imprecise because of small groups, there are some general patterns. There is a negative association between full biological birth order and school points, and the coefficients from each combination are similar. This finding holds across both samples and all combinations, as shown in panels b and d of Fig. 1.⁴ However, the association between overall birth order and school achievement differs between the two samples.

³ The slight differences between the estimates for full siblings and all siblings in certain overlapping combinations between the models in panels A and B of Table A1 on the one hand and panels A and B of Table A2 on the other is due to differences in the control variables when considering full siblings versus all siblings. The results are identical in a simple bivariate model. This does not affect our overall interpretations.

⁴ We also pooled the different combinations and estimated interactions between the birth order coefficients and each combination (results available upon request). The results are fully in line with those reported in Fig. 1 and panel B in both Table A1 and Table A2, with no statistically significant variation across the combinations.

Table 2 Descriptive statistics for the father-based and mother-based samples

Variables	Father-Based Sample		Mother-Based Sample	
	N/Mean	SD/%	N/Mean	SD/%
Outcome Variable				
Grade average (mean, SD)	39.00	8.9	37.96	8.87
Birth Order (N, %)				
Overall birth order				
1	6,476	31.17	6,152	31.53
2	6,476	31.17	6,152	31.53
3	6,476	31.17	6,152	31.53
4	1,357	6.53	1,053	5.40
Full biological birth order				
1	13,171	63.39	12,444	63.79
2	6,872	33.07	6,502	33.33
3	736	3.54	563	2.89
Combinations of Number of Children and Partners (N, %)				
1-2	9,576	46.08	9,276	47.55
2-1	5,736	27.60	6,339	32.49
1-3	1,990	9.58	1,173	6.01
2-2	1,452	6.99	1,176	6.03
3-1	842	4.05	737	3.78
Other	1,183	5.70	808	4.14
Control Variables (mean, SD)				
Year of birth	1991.8	(4.00)	1991.8	(4.07)
Mother's age at birth	26.02	(4.98)	26.04	(4.93)
Father's age at birth	29.14	(5.44)	29.21	(6.04)
Number of Observations	20,776		19,508	

Source: The authors' own calculations based on registry data

For children in the mother-based sample, having a higher overall birth order is negatively associated with educational achievement, as indicated in panel c of Fig. 1.⁵ However, the differences between the last sibling of the first set and the first child of the second set seem to be somewhat less marked in the 2-1 and 2-2 combinations. As previously discussed, a negative effect of union dissolution and related factors will typically affect all children of the first set negatively, which offsets at least some of the negative effects typically associated with having a higher birth order. For children born to mothers who display multipartnered fertility, the negative effects of birth order generally seem to trump these offsetting factors. Table A8 (online appendix) shows that this finding holds across the combinations for the full sample when the schooling achievements of the youngest sibling in the first set are

⁵ An analysis with all combinations pooled and an interaction term between each combination and each birth order coefficient indicates no statistically detectable differences between the different combinations (results available upon request). This is in line with the results reported in Fig. 1 and panel A of Table A1.

Fig. 1 The association between school point average and the two types of birth order by different combinations of number of children per co-parent. The values for the different subgroups are indicated as follows: 1-2 indicates that a person has parented one child with the first co-parent, has parented two children with the second co-parent, and does not have any more children with a potential third co-parent; 2-2 indicates two children with each of two co-parents; and so on. We include analyses for subgroups with more than 800 observations, which is the number of observations needed to detect birth order effects in the nuclear family sample and explains why the 3-1 combination is included for only the father-based sample. The groups included constitute more than 90% of the full samples. All models include controls for fathers' and mothers' age at birth, own cohort, and gender. The figure displays 95% confidence intervals. *Source:* The authors' own calculations based on registry data.

compared with the oldest sibling in the second set and the variation in this difference is examined across the combinations of partners and children. The main finding is therefore a clear negative association between both types of birth order and educational achievement in the mother-based sample. In other words, all children achieve fewer school points on average compared with older half- and full siblings in this sample.

In the father-based sample, the general association between overall birth order and school point average is less clear, as seen in panel a of Fig. 1. In all combinations except 3-1, the oldest child of the second set tends to perform *better* in school when compared with the youngest sibling of the first set. This is most pronounced in the 1-2, 1-3, and 2-2 combinations, for which the coefficients for being second-born (1-2, 1-3) and third-born (2-2) are actually positive. Table A9 (online appendix) provides an estimate of this difference for the whole sample, and it generally holds across the combinations. Again, this is consistent with negative effects of union dissolution lowering the average achievement of the child/children of the first set. On the other hand, the associations between full biological birth order and educational achievement are consistently negative, as indicated in panel b of Fig. 1. Logically, this association can also be seen in panel a given that full biological birth order is a subset of overall birth order; later-born children *in the same set* tend to perform worse compared with their older full siblings. This finding implies that children in this sample typically perform on par with or better than the closest half-sibling but worse than full siblings.

Figure 2 shows the associations between school achievement and overall and full biological birth order effects when all combinations of children per partner are pooled (with full results shown in Tables A3 and A4 in the online appendix). This enables us to investigate overall birth order patterns with higher statistical precision. The figure also includes results from the nuclear family sample for comparison (see Table A5 for full results).

Panels a, b, d, and e of Fig. 2 provide evidence for negative associations between the school point average and overall/full biological birth orders relative to older siblings with the same father (top row) and mother (bottom row). The association between school points and overall birth order is weaker in the father-based sample compared with that found in the nuclear family sample. This is consistent with the findings presented in panel a of Fig. 1, which indicate that negative effects of overall birth order cannot be clearly discerned for children with fathers displaying multipartnered fertility. The remaining negative association between *overall* birth order and school achievement in panel a of Fig. 2 is therefore likely to be driven by the negative association between school points and *full biological* birth order shown in panel b of Fig. 1. This is because of the substantial overlap between the full biological and overall birth order: overall birth order sequences are always composed of subsets of full biological birth orders. The question is therefore whether negative birth order effects operate mainly *across* or *within* sets of children depending on the sex of the parent displaying multipartnered fertility. Our third model takes both types of birth order into account simultaneously, thus providing additional evidence on this issue.

The results from Model 3 are presented in panels c and f of Fig. 2. Panel c indicates that when both types of birth order are included in the models, only the full biological birth order relative to siblings with the same *father* remains negatively associated with educational success. Per contra, the coefficient for the overall birth order is now small and not statistically significant. In the mother-based sample (panel f of Fig. 2), the opposite is true. Here, the overall birth order still matters, whereas the full biological birth order is no longer substantially or statistically significant. These arguably represent estimates of net effects of birth order in the context of multipartnered fertility given that we cannot isolate birth order effects from those of other factors, such as union dissolutions and family stability, using this model.

Fig. 2 The differences in school point average at the end of compulsory education based on two types of birth order, with the firstborn^a as the reference in each case. Coefficients are shown for the two main samples as well as the nuclear family sample, where children with half-siblings are excluded (all panels). All coefficients indicate the difference in the average school points relative to the firstborn for the two types of birth order. Panels a–c display the results from Models 1–3 for the father-based sample. Panels d–f display the results from Models 1–3 for the mother-based sample. All results are based on fixed-effects regression within fathers/mothers (overall birth order) or their combination (full biological birth order). All models include controls for age of father at birth, age of mother at birth, and year of birth. The figure displays 95% confidence intervals. The reference groups for full biological and overall birth order are not identical, as specified in the variable definitions; they are overlapping only within the first set. *Source:* The authors' own calculations based on registry data.

The negative associations between birth order and school achievement documented here are both statistically and substantially significant, and they are likely to have consequences for later school transitions. The largest disadvantages are found in the mother-based sample, where the difference between overall firstborns and second-borns is approximately 2 school points (roughly one-fifth of a standard deviation),

whereas the difference between firstborns and fourth-borns is about 4 school points (roughly one-half of a standard deviation). The difference between first- and fourth-borns is of a magnitude similar to the sex gap in our sample.

Discussion

Summary of the Main Findings and Evaluation of Our Empirical Expectations

Previous studies on birth order and socioeconomic outcomes have typically studied nuclear families. However, recent demographic developments have led to increasingly complex family constellations. We acknowledge the importance of such developments by studying birth order effects on educational achievement in the context of multipartnered fertility. Based on the resource dilution and confluence theories as well as previous demographic research, we derive expectations for each type of birth order for each sample according to the sex of the parent who displayed multipartnered fertility.

We ask three research questions. Our first question asks whether associations between birth order and educational achievement exist in the context of multipartnered fertility. We provide evidence for overall negative associations between both full biological and overall birth order on the school point average at the end of compulsory schooling in both samples.

Our second and third research questions ask whether this association differs for full biological and overall birth order or according to the sex of the parent displaying multipartnered fertility. The answers to these two questions prove to be interrelated. In the mother-based sample, the association between overall and full biological birth order with educational achievement was negative. In the father-based sample, the overall birth order effect is weaker and less clear-cut, whereas the negative effect of full biological birth order remains important. Furthermore, when both overall and full biological birth order are included in the same model, only the overall birth order remains negatively associated with school achievements in the mother-based sample. For siblings in the father-based sample, this is the other way around, with only the full biological birth order remaining significant. We consider these associations as credible net effects of negative birth order effects due to the impact of union dissolution and family complexity, which clearly have a negative impact on school achievement.

We interpret these findings as indicating that the negative effects of birth order on educational achievement mainly operate *within* sets of children who share the same father and *across* sets for siblings born to the same mother. Asking whether one's birth order relative to all siblings is more important for upper secondary school achievement than one's birth order relative to full siblings yields a more intuitive interpretation. For siblings born to mothers displaying multipartnered fertility, the overall birth order across sets of siblings is clearly the most important: on average, these children perform worse in school compared with their older full and half-siblings, and the difference is similar in magnitude to that found in nuclear families.

For siblings born to fathers who display multipartnered fertility, the full biological birth order is what matters the most: these children tend to obtain fewer school points, on average, compared with older *full siblings* but also tend to perform equally well or

better compared with the closest older *half-sibling*. The latter finding is likely due to factors other than typical birth order mechanisms, such as union dissolutions, family complexity, and differences in union stability. Although these mechanisms are relevant in both samples, the evidence indicates that they are not trumped by negative overall birth order effects in the mother sample. Still, there is clearly room for variation across the different subgroups, and future research should investigate this further.

Our findings are consistent with previous research. In a supplementary analysis, Kristensen and Bjerkedal (2010) considered birth order effects for mothers regardless of the number of co-parents and provided evidence for negative birth order effects on educational attainment at age 25. Other studies included control variables for union dissolution in the nuclear family birth order design, which did not make a difference (Kantarevic and Mechoulan 2006; Sigle-Rushton et al. 2014). However, the last two studies used two-parent fixed effects, and neither of the three took the issue of multipartnered fertility explicitly into account.

Our evidence supports the theoretical expectations that we derive from the resource dilution and confluence theories. In our expectations, we distinguish between the behavior of mothers and fathers who display multipartnered fertility. We argue that children who share a mother displaying multipartnered fertility will face a more consistent environment, where birth order effects can arise across sets. Fathers, on the other hand, are less likely to live with and interact with all their children at the same time for an extended period. This implies that although the father is often present in his children's life for some time, we will not see the same consistency across sets in the father sample, thereby making full biological birth order effects more relevant. Based on the two theories, the question is whether the resource dilution and intellectual climate are mostly relevant only for full siblings in the father-based sample or both half- and full siblings in the mother-based sample. Our findings are consistent with these expectations. The confluence and resource dilution theories are therefore able to account for birth order patterns outside a nuclear family context. Still, our results indicate that they need modification or supplementation and need to take parental sex and related behavioral differences into account.

Alternative Interpretations

Still, other theories and mechanisms could also provide relevant explanations. A more radical interpretation is that mechanisms specific to the mother are more important in explaining the birth order effects we observe, and possibly also birth order effects documented in other contexts.⁶ Several mechanisms are plausible based on the existing literature. For example, mothers in Western countries typically drink and smoke more during later pregnancies (Lundberg et al. 2010), and mothers breastfeed their first children more than their later-born children and are more likely to seek prenatal care for the oldest (Buckles and Kolka 2014). This interpretation also aligns with recent evidence on cognitive test scores from age 1 until school age using direct measures of

⁶ We are here referring to the negative associations between birth order and educational achievements. To explain processes relating to union dissolutions, family stability, and so forth, other factors and the father's role must clearly be taken into account, as previously discussed.

maternal and paternal involvement. Only maternal involvement proved relevant in explaining negative birth order effects (Lehmann et al. 2018).

This mother-centric interpretation could favor the resource dilution theory, which can easily be modified to ascribe a greater importance to maternal resources and investments. In terms of the confluence theory, this interpretation would mean that the mother has a more positive influence on the average intellectual climate of the family compared with the father. This is hard to square with the fact that the confluence theory is based on the distinction between the intellectual levels of children and adults. Additionally, controlling for spacing does not affect our overall findings (see Tables A10 and A11 in the online appendix). Because spacing effects are emphasized in the confluence theory (e.g., Zajonc et al. 1979), this finding could indicate that mechanisms associated with the resource dilution theory are more important. Still, further data are necessary to separate the predictions of the two theories on a more fine-grained level, and evidence with direct measures of paternal and maternal quality time is mixed (Lehmann et al. 2018; Monfardini and See 2016).

It would be possible to adjudicate between our main explanation and the mother-centric interpretation if paternal investments could be measured directly given that the time spent with the father should not be driving birth order patterns if the first explanation is correct. However, as emphasized previously, we are unable to measure parental presence and investments nor the existence of shared residence arrangements directly. In a set of additional analyses with reduced samples, we included information on the number of years in which fathers lived with both the mother and their common children. We included information on years of cohabitation for each individual child as well as the average number of years in each set. This did not affect the main results appreciably (see Tables A12 and A13 in the online appendix). Still, this is a rough proxy for paternal involvement, and some fathers are more involved with nonresident children than others, while a very small group has full physical custody (Lyngstad et al. 2014).⁷

A third possibility, of course, is that other differences between the two samples are driving the findings. For example, perhaps the likelihood of the co-parents themselves displaying co-partner fertility could lead to differences between sets that are confounded with birth order mechanisms, for example, based on sex differences in the propensity for displaying this behavioral pattern. We investigated the relevance of this explanation by performing separate analyses depending on whether the first, second, or either co-parent themselves displayed multipartnered fertility (see Tables A14 and A15 in the online appendix). Despite differences in the magnitude of the estimates in the mother sample, the general pattern of full biological and overall birth order effects and mother/father sample remained: the differences were not statistically significant, and some subgroups were very small, thereby increasing statistical uncertainty.

Sensitivity Checks and Remaining Limitations

Several analyses were also carried out to evaluate the robustness of our findings and assumptions. We performed two sensitivity analyses with slightly reduced samples in a

⁷ Although cohabitation data enable us to identify something close to fathers with sole custody, the number of individuals was too small for meaningful statistical analysis.

further attempt to separate the negative consequences of union dissolutions from those of birth order.⁸ The first shows that accounting for union dissolutions does not affect the estimate for the nuclear family sample (Table A16, online appendix), indicating that the occurrence of union dissolutions does not affect our estimate of full biological birth order effects. In the second analysis, we considered only cases in which *both* unions of the parent displaying multipartnered fertility were dissolved in our two main samples, thus exposing both sets of children to this negative influence (Tables A17 and A18, online appendix). The findings were very similar to those of our main results, which indicates that the size of the downward bias in birth order effects caused by union dissolutions and related processes are not large enough to be detected given our sample size and statistical uncertainty in our main samples.

Our inclusion criteria require that parents have several children with multiple partners in a relatively short span of time (maximally 13 years between the first and the last), which could affect the external validity of our findings. However, previous research on multipartnered fertility has indicated that this fertility pacing is quite common for this group (Lappegård and Rønsen 2013). Also, previous studies have indicated that individuals displaying multipartnered fertility typically differ in their level of education from the rest of the population (Carlson and Furstenberg 2006; Evenhouse and Reilly 2016; Guzzo and Furstenberg 2007; Lappegård and Rønsen 2013; Thomson et al. 2014; Wiik and Dommermuth 2014). In four sensitivity analyses, looking at both our main samples and the nuclear family sample, we did not find that birth order effects vary in a statistically meaningful way based on either parent's educational level (see Tables A19–A22, online appendix). This finding is in line with previous research on nuclear families (Black et al. 2005; Härkönen 2014). Therefore, we do not expect the magnitude of birth order effects documented in our study to be affected by the selection in our sample, at least not based on observable educational characteristics.

Conclusion

We provide novel evidence for negative birth order effects that vary depending on the sex of the parent displaying multipartnered fertility. This type of fertility pattern has become increasingly common in western countries, and is likely to become even more common with time because of the continuation of relevant demographic trends and because of the likely transmission of the behavior across generations (Lappegård and Thomson 2018). Our main interpretation is that the pattern we document results from a combination of birth order mechanisms and residential arrangements. This makes our findings relevant for policy-makers given that the prevalence of different residential arrangements is influenced by both social norms and family policies (Cancian et al. 2014; Fehllberg et al. 2011; Sodermans et al. 2013). Shared residence between mothers and fathers has become more common in Norway and several other countries, and 50/50 shared residence has become a more important normative and political goal (Harris-Short 2010; Lyngstad et al. 2014; Singer 2008). This increase was particularly marked

⁸ These samples include siblings in the 1987–1998 cohort. Information on cohabitation is available for this period only.

during the 2000s, whereas the youngest children in this study were born in 1998. During the 1980s and mid-1990s, shared residence was extremely rare in Norway (Jensen and Clausen 1997). Future studies should use data that include later cohorts or investigate countries with different family policies. Although such descriptive evidence is itself valuable, this could also enable adjudication between our suggestions that residence arrangements, the role of mothers more generally, or other potential mechanisms explain our birth order findings. Such research may provide new and important evidence to better understand the genesis of birth order effects, both in this context and more generally.

Acknowledgments Open Access funding provided by University of Oslo (including Oslo University Hospital). The authors are grateful for comments from five anonymous reviewers and the editorial board of *Demography*; Gunn Elisabeth Birkelund; Trude Lappegård; Torkild Hovde Lyngstad; participants in the Segregation seminar at the university of Oslo; and participants in the Demographic Breakfast Seminar at OsloMet – Oslo Metropolitan University.

Authors' Contributions Both authors have contributed to the conceptualization, data setup, analyses, writing (reviewing and editing), and visualization. Lillehagen wrote the first draft and developed the original idea.

Data Availability The data required for this project were made available through the SEGREGATION project funded by the Research Council of Norway (RCN project #202479). These registry data are available from Statistics Norway to researchers with projects that satisfy the data owners' requirements. For information on how to gain access to Norwegian microdata and formal requirements, see www.ssb.no/en/omssb/tjenester-og-verktoy/data-til-forskning.

Compliance With Ethical Standards

Conflict of Interest The authors have declared that no competing interests exist.

Ethics and Consent Access to administrative register data was approved by The Norwegian National Committee for Research Ethics in the Social Sciences and the Humanities (NESH), the NSD Data Protection Services (Personvernombudet), and Statistics Norway. These data are de-personalized and range the full population.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andersson, G. (2004). Childbearing developments in Denmark, Norway, and Sweden from the 1970s to the 1990s: A comparison. *Demographic Research*, 3, 155–176. <https://doi.org/10.4054/DemRes.2004.S3.7>
- Andersson, G., & Kolk, M. (2015). Trends in childbearing, marriage and divorce in Sweden: An update with data up to 2012. *Finnish Yearbook of Population Research*, 50, 21–30.

- Bakken, A., & Elstad, J. I. (2012). *For store forventninger? Kunnskapsløftet og ulikhetene i grunnskolekarakterer* [Expectations set too high? Changes in school policy and inequalities in compulsory school grades] (NOVA Report, Vol. 7/12). Oslo, Norway: Norsk Institutt for Forskning om Oppvekst, Velferd og Aldring.
- Barclay, K., Hällsten, M., & Myrskylä, M. (2017). Birth order and college major in Sweden. *Social Forces*, 96, 629–660.
- Barclay, K., & Myrskylä, M. (2014). Birth order and physical fitness in early adulthood: Evidence from Swedish military conscription data. *Social Science & Medicine*, 123, 141–148.
- Barclay, K. J. (2015). Birth order and educational attainment: Evidence from fully adopted sibling groups. *Intelligence*, 48, 109–122.
- Barclay, K. J., & Kolk, M. (2018). Birth intervals and health in adulthood: A comparison of siblings using Swedish register data. *Demography*, 55, 929–955.
- Baxter, J. A. (2012). Fathering across families: How father involvement varies for children when their father has children living elsewhere. *Journal of Family Studies*, 18, 187–201.
- Beaujouan, É. (2016). Second unions now more stable than first? A comparison of separation risks by union order in France. *European Journal of Population*, 32, 293–321.
- Berger, L. M., Cancian, M., & Meyer, D. R. (2012). Maternal re-partnering and new-partner fertility: Associations with nonresident father investments in children. *Children and Youth Services Review*, 34, 426–436.
- Bernardi, F., & Radl, J. (2014). The long-term consequences of parental divorce for children's educational attainment. *Demographic Research*, 30, 1653–1680. <https://doi.org/10.4054/DemRes.2014.30.61>
- Björklund, A., Eriksson, T., Jäntti, M., Raaum, O., & Österbacka, E. (2004). Family structure and labor market success: The influence of siblings and birth order on the earnings of young adults in Norway, Finland, and Sweden. In M. Corak (Ed.), *Generational income mobility in North America and Europe* (pp. 208–225). Cambridge, UK: Cambridge University Press.
- Black, S. E., Devereux, P. J., & 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, 669–700.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2011). Older and wiser? Birth order and IQ of young men. *CEifo Economic Studies*, 57, 103–120.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2016). Healthy(?), wealthy, and wise: Birth order and adult health. *Economics & Human Biology*, 23, 27–45.
- Bonesrønning, H., & Massih, S. S. (2011). Birth order effects on young students' academic achievement. *Journal of Socio-Economics*, 40, 824–832.
- Boomsma, D. I., van Beijsterveld, T. C., Beem, A. L., Hoekstra, R. A., Polderman, T. J., & Bartels, M. (2008). Intelligence and birth order in boys and girls. *Intelligence*, 36, 630–634.
- Brenøe, A., & Molitor, R. (2015). *Birth order and health of newborns: What can we learn from Danish registry data?* (SSRN working paper). Retrieved from <https://ssrn.com/abstract=2665093>
- Bronte-Tinkew, J., Horowitz, A., & Scott, M. E. (2009). Fathering with multiple partners: Links to children's well-being in early childhood. *Journal of Marriage and Family*, 71, 608–631.
- Buckles, K., & Kolka, S. (2014). Prenatal investments, breastfeeding, and birth order. *Social Science & Medicine*, 118, 66–70.
- Butikofer, A., Løken, K. V., & Salvanes, K. G. (2015). *Long-term consequences of access to well-child visits* (NHH Department of Economics Discussion Paper No. 29/2015). Retrieved from <https://doi.org/10.2139/ssrn.2719404>
- Cancian, M., & Meyer, D. R. (1998). Who gets custody? *Demography*, 35, 147–157.
- Cancian, M., Meyer, D. R., Brown, P. R., & Cook, S. T. (2014). Who gets custody now? Dramatic changes in children's living arrangements after divorce. *Demography*, 51, 1381–1396.
- Cancian, M., Meyer, D. R., & Cook, S. T. (2011). The evolution of family complexity from the perspective of nonmarital children. *Demography*, 48, 957–982.
- Carlson, M. J., & Furstenberg, F. F. (2006). The prevalence and correlates of multipartnered fertility among urban U.S. parents. *Journal of Marriage and Family*, 68, 718–732.
- Carneiro, P., Løken, K. V., & Salvanes, K. G. (2015). A flying start? Maternity leave benefits and long-run outcomes of children. *Journal of Political Economy*, 123, 365–412.
- Conley, D. (2004). *The pecking order: Which siblings succeed and why*. New York, NY: Pantheon Books.
- De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, 29, 576–588.
- Dewey, K. G., & Cohen, R. J. (2007). Does birth spacing affect maternal or child nutritional status? A systematic literature review. *Maternal & Child Nutrition*, 3, 151–173.

- Downey, D. B. (2001). Number of siblings and intellectual development: The resource dilution explanation. *American Psychologist*, 56, 497–504.
- Evenhouse, E., & Reilly, S. (2016). Multipartnered fertility. In C. L. Shehan (Ed.), *Encyclopedia of family studies*. Wiley Blackwell: Hoboken, NJ. <https://doi.org/10.1002/9781119085621.wbef376>
- Fehlberg, B., Smyth, B., Maclean, M., & Roberts, C. (2011). *Caring for children after parental separation: Would legislation for shared parenting help children?* (Family Policy Briefing No. 7). Oxford, UK: University of Oxford.
- Fomby, P., Goode, J. A., & Mollborn, S. (2016). Family complexity, siblings, and children's aggressive behavior at school entry. *Demography*, 53, 1–26.
- Fomby, P., & Osborne, C. (2017). Family instability, multipartner fertility, and behavior in middle childhood. *Journal of Marriage and Family*, 79, 75–93.
- Foster, J. W., & Archer, S. J. (1979). Birth-order and intelligence: Immunological interpretation. *Perceptual and Motor Skills*, 48, 79–93.
- Grinberg, A. (2015). The effect of birth order on occupational choice. *Atlantic Economic Journal*, 43, 463–476.
- Guzzo, K. B. (2014). New partners, more kids: Multiple-partner fertility in the United States. *Annals of the American Academy of Political and Social Science*, 654, 66–86.
- Guzzo, K. B., & Furstenberg, F. F. (2007). Multipartnered fertility among American men. *Demography*, 44, 583–601.
- Härkönen, J. (2014). Birth order effects on educational attainment and educational transitions in West Germany. *European Sociological Review*, 30, 166–179.
- Harris-Short, S. (2010). Resisting the march towards 50/50 shared residence: Rights, welfare and equality in post-separation families. *Journal of Social Welfare and Family Law*, 32, 257–274.
- Hertwig, R., Davis, J. N., & Sulloway, F. J. (2002). Parental investment: How an equity motive can produce inequality. *Psychological Bulletin*, 128, 728–745.
- Jensen, A.-M., & Clausen, S.-E. (1997). *Samvær og fravær: Foreldres kontakt med barn de ikke bor sammen med* [Being together and being apart: Parental contact with nonresident children] (Research Note 1997: 103). Oslo, Norway: Norsk Institutt for by- og Regionforskning.
- Kalmijn, M. (2015a). Family disruption and intergenerational reproduction: Comparing the influences of married parents, divorced parents, and stepparents. *Demography*, 52, 811–833.
- Kalmijn, M. (2015b). Father-child relations after divorce in four European countries: Patterns and determinants. *Comparative Population Studies*, 40, 251–276.
- Kantarevic, J., & Mechoulam, S. (2006). Birth order, educational attainment, and earnings: An investigation using the PSID. *Journal of Human Resources*, 41, 755–777.
- Kitterød, R. H., Lidén, H., Lyngstad, J., & Wiik, K. A. (2016). *Delt bosted for barna etter samlivsbrudd – en praksis for folk flest?* [Shared living arrangements for children after a union dissolution – A common practice?]. *Sosiologisk Tidsskrift*, 24, 27–50.
- Kitterød, R. H., & Lyngstad, J. (2012). Untraditional caring arrangements among parents living apart: The case of Norway. *Demographic Research*, 27, 121–152. <https://doi.org/10.4054/DemRes.2012.27.5>
- Kristensen, P., & Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, 316, 1717.
- Kristensen, P., & Bjerkedal, T. (2010). Educational attainment of 25 year old Norwegians according to birth order and gender. *Intelligence*, 38, 123–136.
- Lappegård, T., & Rønsen, M. (2013). Socioeconomic differences in multipartner fertility among Norwegian men. *Demography*, 50, 1135–1153.
- Lappegård, T., Rønsen, M., & Skrede, K. (2011). Fatherhood and fertility. *Fathering*, 9, 103–120.
- Lappegård, T., & Thomson, E. (2018). Intergenerational transmission of multipartner fertility. *Demography*, 55, 2205–2228.
- Lehmann, J. Y. K., Nuevo-Chiquero, A., & Vidal-Fernandez, M. (2018). The early origins of birth order differences in children's outcomes and parental behavior. *Journal of Human Resources*, 53, 123–156.
- Lesthaeghe, R. (2010). The unfolding story of the second demographic transition. *Population and Development Review*, 36, 211–251.
- Lichter, D. T., Turner, R. N., & Sassler, S. (2010). National estimates of the rise in serial cohabitation. *Social Science Research*, 39, 754–765.
- Lundberg, F., Cnattingius, S., D'Onofrio, B., Altman, D., Lambe, M., Hultman, C., & Iliadou, A. (2010). Maternal smoking during pregnancy and intellectual performance in young adult Swedish male offspring. *Paediatric and Perinatal Epidemiology*, 24, 79–87.
- Lyngstad, J., Kitterød, R. H., & Nymoen, E. H. (2014). *Bosted og samvær 2002, 2004 og 2012: Endringer i ansvar og omsorg for barna når mor og far bor hver for seg* [Place of residence and custody 2002, 2004

- and 2012: Changes in responsibility and care for children when mothers and fathers live apart] (Report 2014, 2). Oslo, Norway: Statistisk Sentralbyrå.
- Manning, W. D., & Smock, P. J. (1999). New families and nonresident father-child visitation. *Social Forces*, 78, 87–116.
- McLanahan, S., Tach, L., & Schneider, D. (2013). The causal effects of father absence. *Annual Review of Sociology*, 39, 399–427.
- Monfardini, C., & See, S. G. (2016). Birth order and child cognitive outcomes: An exploration of the parental time mechanism. *Education Economics*, 24, 481–495.
- Morgan, S. L., & Winship, C. (2015). *Counterfactuals and causal inference* (2nd ed.). New York, NY: Cambridge University Press.
- Pavan, R. (2016). On the production of skills and the birth order effect. *Journal of Human Resources*, 51, 699–726. <https://doi.org/10.3368/jhr.51.3.0913-5920R>
- Petren, R. E. (2017). Paternal multiple partner fertility and environmental chaos among unmarried nonresident fathers. *Journal of Social Service Research*, 43, 100–114.
- Poortman, A. R., & Lyngstad, T. H. (2007). Dissolution risks in first and higher order marital and cohabiting unions. *Social Science Research*, 36, 1431–1446.
- Price, J. (2008). Parent-child quality time: Does birth order matter? *Journal of Human Resources*, 43, 240–265.
- Sigle-Rushton, W., Lyngstad, T. H., Andersen, P. L., & Kravdal, Ø. (2014). Proceed with caution? Parents' union dissolution and children's educational achievement. *Journal of Marriage and Family*, 76, 161–174.
- Silles, M. A. (2010). The implications of family size and birth order for test scores and behavioral development. *Economics of Education Review*, 29, 795–803.
- Singer, A. (2008). Active parenting or Solomon's justice? Alternating residence in Sweden for children with separated parents. *Utrecht Law Review*, 4(2), 35–47.
- Skjorten, K., Barlundhaug, R., & Lidén, H. (2007). *Delt bosted for barn* [Shared residence for children]. Oslo, Norway: Gyldendal Akademisk.
- Sodermans, A. K., Matthijs, K., & Swicegood, G. (2013). Characteristics of joint physical custody families in Flanders. *Demographic Research*, 28, 821–848. <https://doi.org/10.4054/DemRes.2013.28.29>
- Steele, F., Sigle-Rushton, W., & Kravdal, Ø. (2009). Consequences of family disruption on children's educational outcomes in Norway. *Demography*, 46, 553–574.
- Sulloway, F. J. (2007). Birth order and intelligence. *Science*, 316, 1711–1712.
- Tach, L., Mincy, R., & Edin, K. (2010). Parenting as a “package deal”: Relationships, fertility, and nonresident father involvement among unmarried parents. *Demography*, 47, 181–204.
- Thomson, E. (2014). Family complexity in Europe. *Annals of the American Academy of Political and Social Science*, 654, 245–258.
- Thomson, E., Lappegård, T., Carlson, M., Evans, A., & Gray, E. (2014). Childbearing across partnerships in Australia, the United States, Norway, and Sweden. *Demography*, 51, 485–508.
- Wiik, K. A., & Dommermuth, L. (2014). Who remains unpartnered by mid-life in Norway? Differentials by gender and education. *Journal of Comparative Family Studies*, 45, 405–424.
- Zajonc, R. B. (1983). Validating the confluence model. *Psychological Bulletin*, 93, 457–480.
- Zajonc, R. B., & Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review*, 82, 74–88.
- Zajonc, R. B., Markus, H., & Markus, G. B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology*, 37, 1325–1341.