Finding the Land of Opportunity

Intergenerational Mobility in Denmark

Jesper Eriksen
Department of Business and Management
Aalborg University

178,866 characters with spaces - 74.5 pages

May 31, 2018

Abstract

Denne afhandling omhandler lige muligheder inden for intergenerationel mobilitet i en dansk kontekst. Afhandlingen forholder sig til spørgsmålet, "Er lighed i muligheder påvirket af lokalområder i Danmark?". Jeg besvarer forskningsspørgsmålet ved at undersøge tre underspørgsmål: (1) "Hvad karakteriserer den geografiske fordeling af muligheder i Danmark?", (2) "Hvorfor kunne bestemte grupper, defineret ved geografi, forventes at opleve højere eller lavere mobilitet? Hvilken rolle spiller lokalområde karakteristika som forklarende variabel i forhold til lige muligheder?", og (3) "Hvad er effekten af at flytte fra ét lokal område til et andet for lige muligheder for flyttende familier?". Først opstiller jeg et teoretisk grundlag for at kunne diskutere lighed i muligheder, som filosofisk begreb, i relation til gængse mål for lighed i muligheder: intergenerationel mobilitet. Jeg påviser, at det ikke er muligt at vurdere, hvorvidt lighed i muligheder påvirkes ud fra nogen af de tre teorier ved alene at undersøge mobilitetsmål. For at besvare forskningsspørgsmålet kræves der information om, hvordan lokalområder kausalt påvirker intergenerationel mobilitet.

For at besvare første underspørgsmål beskriver jeg grundlæggende træk i litteraturen om intergenerational mobilitet, samt den metodiske literatur om intergenerationel mobilitetsmål. Datagrundlaget for besvarelsen er dansk registerdata for perioden 1980-2012, som tillader kobling af børn og forældre, samt måling af samlet indkomst før skat. Ved at allokere børn født mellem 1973-1975 til de kommuner, som primært vokser op i, fra de er 7-15 viser jeg, at der er substantiel variation i intergenerationel mobilitet på tværs af danske kommuner. Med udgangspunkt i diskussionen af teorierne om lighed i muligheder, er det ikke muligt at vurdere om lighed i mulighed faktisk varierer fra kommune til kommune. Resultatet kan alene indikerer, at det kan være er tilfældet. Herefter introducerer jeg "Neighborhood Effects" litteraturen, der omfatter en række teoretiske bud på, hvorfor variation i karakteristika for lokalområder skulle påvirke lighed i muligheder. Med udgangspunkt i teorierne opstiller jeg 9 forskellige variable som indikatorer for lokalområde literaturen; andelen af arbejdsløse til befolkningen i alderen 15-64, andelen af individer, som tilhører den national tiende indkomst

decil, andelen af individer, som tilhører den nationale nederste indkomst decil, andelen af teenagemødre til befolkningen over 15, andelen af teenagefødsler, andelen af første og andengenerations ikke-vestlige immigranter, andelen af højtuddannede, gini koefficienten som mål for økonomisk ulighed, og antallet af rapporterede kriminelle handlinger relativt til befolkningsstørrelsen. Jeg finder, at målene ikke ko-varierer med mål for relativ mobilitet, men at de ko-varierermed absolut mobilitet, den forventede indkomstrang for et barn født af forældre ved den 25. indkomst percentil i forældrenes indkomst fordeling. Fundet indikerer altså, at mobilitet er relateret til karakteristika for kommuner og deres indbyggere. Dog kan resultatet ikke afgøre, hvorvidt der faktisk er substantielle forskelle i lige muligheder mellem kommuner. Den observerede ko-variation kan være drevet af selv-selektering. Tredje underspørgsmål besvares med henblik på at vurdere, om der er kausale effekter på intergenerational mobilitet af lokalområder, approksimeret af kommuner. Her opstiller jeg et kausal estimations framework, som undersøger tilpasningen i forventet indkomstrang for to sæt af børn, som flytter mellem kommuner. Den første stikprøve består af børn, som flytter én gang i alderen 8-15, den anden af børn, der flytter én gang i alderen 8-20. På baggrund af parametriske og semi-parametriske tests afviser jeg, at lokalområder har en kausal effekt på intergenerational mobilitet på kommunalt niveau.

Konklusionen på afhandlingen er, at der ikke er tegn på at lighed i muligheder er påvirket af lokalområder i Danmark, når lokalområder er opgjort som kommuner.

Contents

Preface

Several acknowledgements are due. It is the culmination of many years of studies, and I remain grateful to the fellow students that I have had the joy to interact with at both Aalborg University and my semester spent abroad. Some, of course deserves particular thanks in excess of what a grateful student might give. In Spring of 2016, I contacted members of the IKE research group at Aalborg University, looking for opportunities to work as a research assistant. To my luck, Ina Drejer found it in her heart to not only hire me, but ask me if I would consider becoming the first 4+4 PhD student at the department of Business and Management. I could hardly say no, and started what has been two years of seemingly endless learning in the warm company of colleagues in the IKE group. The position also came with the opportunity to work on subjects that I have come to find near my heart: economics of education and social mobility. These interests are at the very foundation of the present thesis. It appears to me that few questions are as important as why we fare the way we do in life - why do some become homeless and left for living in city streets, while others good to elite colleges and start careers with prospects of security and satisfaction for years to come? Are there substantial differences in society or within us that drives these differences, or is it simply a game of chance which cannot be predicted. At is core, these questions touches on the notion of Equality of Opportunity, a hallmark idea of modern as well as ancient philosophical thoughts on how to create the best society for its citizens. Delving into the work of this thesis, I have had the opportunity to not only learn from the best, but develop my thinking substantially on these questions and ideas.

Throughout the process of twisting and turning data, reading and writing endless pages, several people have shown their good will and listened provided comments as I rambled about intergenerational elasticities and correlations, and Rawls notion of fair equality of opportunity. Among them are Søren Etzerodt, a friend, colleague, and co-author, who has been a part of inspiring and sparking my interest in doing research since our very first shared semester at Aalborg University. Having started our research journeys together, he has also, mercifully, read and commented on substantial parts of the thesis. Maria Junker Nielsen has spent hours combing through my findings with seemingly endless enthusiasm, as well as been a solid friend through years of studies in Aalborg. Similarly, I have shared many a pleasant conversation with Jonas Thorborg Stage as we together deciphered the deeper questions of the thesis. Frederik Treney deserves thanks, not only for helping me with the editing of sections, but also for the constantly good mood that he has helped bring about even when writing was tough. Most dearly to me, is Kat Stolerman. She has been a constant source of joy and and a steadying force throughout the last months of writing. She is the primary source of all good writing skills that I have developed (and hopefully employed) for this thesis. She has saved many a drafted section from atrocities towards the gods of grammar, and deserves credit for all readability of the thesis.

Naturally, a special thanks also goes to my two PhD supervisors, Jacob Rubæk Holm, and and Martin David Munk. For the last two years, Jacob has been a role model of how to deal with the world of academia, and a steadfast supporter of all of the (sensible) ideas that I have put before him throughout the last two years. As one walks through endless days of work, the knowledge that your supervisor has your back is one of the strongest motivators to keep pushing when outcomes are unknown. Martin has not only been a PhD supervisor, but also the supervisor for this thesis. From many conversations, I have developed my understanding of the field of intergenerational mobility, and without his guidance, needless to say, this thesis would not have come to its present standing were it not for Martin's comments.

It remains only to be said that all errors are my own and that I hope you, dear reader, will find the content of this thesis as interesting as I have while writing it.

Jesper Eriksen

Aalborg, Denmark

May 2018

Chapter 1

Introduction

Equality of opportunity is a hallmark idea of most western societies. It is embodied in the infamous "American Dream", and a foundational and embraced idea across the political spectrum. While often referred to as a societal goal, much research has shown that equality of opportunity may in fact be only just a goal. In the intergenerational mobility literature, present in both economics and sociology, it has been shown that a child's parental background continues to be a strong predictor of income, and educational and occupational achievement.The same is true for the neighborhoods that children grow up in. Neighborhood effects research has shown that growing up in a deprived area can with substantial precision predict the outcomes of the child, as well as the future areas the child will live in. While the goal of equality of opportunity is broadly agreed upon, it appears that it has not yet been achieved.

The research on intergenerational mobility, which investigate the extend of equality of opportunity, has seen a significant increase in interest within the last 20 years, partly owing to increased availability of multi-generational sample and administrative data, and to research that has shown rising inequalities in income and wealth over time (Atkinson et al., 2011; Piketty, 2014). There are few signs that the interest in the topic is decreasing - in 2017, for example, a complete issue in the Scandinavian Journal of Economics was devoted to the topic (Durlauf and Mogstad, 2017). Several sources provide valuable overviews of national-level mobility estimates, including Blanden (2013), Corak (2006a), Corak (2016), and Torche

(2015). The main finding in the literature has been that there is a strong statistical relationship between parents and children's outcomes (in income, education, and occupation) in particular in the US and UK. In a Danish context income mobility studies remain relatively scarce and tend to emphasize national level statistics (see e.g. Bratsberg et al. 2007a, Boserup et al. 2016, Hussain et al. 2009, Landersø and Heckman 2017, and Munk et al. 2016). The findings that are available for Denmark and other Scandinavian countries (e.g. Jäntti et al. 2006; Corak 2004) suggest that these nordic societies are characterised by substantially less persistence in outcomes between parents and children. They are, it seems, better at providing equality of opportunity. But due to the fact that most of the literature revolves around nationally comparative studies and methodological issues, little is known about how mobility varies within countries, in particular Scandinavian countries such as Denmark.

A recent paper by Chetty et al. (2014) documents large differences in US intergenerational mobility across commuting zones. These differences in mobility correlate with local area characteristics. For instance, residential segregation and income inequality is negatively correlated with upward mobility measured as the expected income rank for a child born to parents in the 25th percentile of the parental income distribution. Other correlates include quality of primary schools, social capital, and family stability in the local area, which all relate positively with mobility. Using these findings, Sharkey and Torrats-Espinosa (2017) show that higher levels of crime predicts lower upward mobility, and Rothwell and Massey (2015a) also finds support for the positive relation with higher quality schooling. An important consideration with correlations is whether they correspond to causal effects; do neighborhoods (or, as in Chetty et al. (2014), commuting zones) matter for the outcomes of children? The correlations can be driven by parents self-selecting into areas that have characteristics that match the expected outcomes their children would otherwise have. Sharkey (2008), for example, shows that among the participants in the Panel Study of Income Dynamics data, the persistence in neighborhood quality (measuring e.g. by poverty) is up to 60 percent in the US. This and related neighborhood research (see e.g. Sampson (2012)) supports a claim of individuals self-selecting into particular local communities for various reasons. In a follow-up paper to the 2014 study, Chetty (2016) shows that there are strong indications of causal effects of neighborhoods at the commuting zone and county level. They find that for children who move to a new area, each additional year spent in a (1 percentage point) better area[1] increases the expected income children (by 0.04 percentage point). These findings of causal effects of neighborhoods are in line with more qualitative work by e.g. Sampson (2012) and Putnam (2015), who show that opportunities of children in the US differ markedly between different areas and income levels due to, among other things, lacking local institutional capacity to support them. The infamous Moving To Opportunity Experiment provides another example. The large scale experiment sent families from high poverty inner city areas to lower poverty with the aid of rent support (Gennetian et al., 2011). While there were no signs of short term improvements of economic outcomes for the families (Ludwig et al., 2008), Chetty et al. (2016a) have found positive effects on income in adulthood for the children that moved to areas with less poverty. Together these findings on the effects of neighborhoods suggest that equality of opportunity is limited by location.

The strong US research agenda on neighborhoods and equality of opportunity leaves open research questions in a Scandinavian context. The intergenerational mobility literature would suggest that the extent of inequality of opportunity in universal welfare societies is less severe than in the US. In this thesis I turn to investigating this contention, focusing on equality of opportunity. As the main research question of the thesis, I ask "Is equality of opportunity affected by neighborhoods in Denmark?" In order to answer the main question, I will answer three sub-questions in turn.

1. What characterizes the spatial distribution of opportunity in Denmark?
2. Why might particular groups defined by geography experience higher or lower degrees of equality of opportunity?
   - What role do neighbourhood characteristics play as predictors of equality of opportunity?
3. What are the effects of changing neighbourhood for equality of opportunity of families?

[1] They measure the difference in area quality by the difference in the expected outcome rank for children, given their parents income rank, asking: Given that a child is born to parents at the 25th percentile income rank in the national income distribution, what is the expected outcome of the child in the area the child moves from, and in the area the child moves to? The difference indicates the difference in absolute mobility between areas.

The first sub-question asks whether if at all there are differences in equality of opportunity within Denmark. The question is exploratory. If we find that there are no spatial differences in equality of opportunity within the country, then there will be no role for neighborhoods to explain differences. The second sub-question asks whether neighborhood characteristics can predict higher and lower equality of opportunity. To answer this, I draw on existing theory and literature to map out factors that can be driving differences in equality of opportunity. To indicate whether these theories and studies may have explanatory power, I investigate whether specific neighborhood related factors have predictive power on the differences in equality of opportunity. The findings are only predictive, and not causal. They cannot tell us whether the neighborhoods themselves, or the self-selection of families into areas with the particular differences is driving the predictive power. The third sub-question therefore takes on the question of causality. Drawing on recent research, I develop an approach, based on the intergenerational mobility models, to infer whether neighborhoods actually affect equality of opportunity.

The answers to the sub-questions provide several contributions to the literature on intergenerational mobility and neighborhood effects. Firstly, I document substantial variation in intergenerational income mobility across municipalities in Denmark using a variety of measures. While there are limited comparisons to draw on, I show that for particular measures Danish spatial distributions of mobility varies by half of that of the US, and mobility is in general higher across Danish municipalities than US commuting zones. Secondly, I show that neighborhood factors do not predict the variation in *relative* mobility - the relation between parents and children's outcomes. Neighborhood factors do, however, predict absolute mobility differences, expected child outcomes for children born to parents at the 25th income percentile. Thirdly, I show that unlike the US case, there is not substantial evidence for persistent neighborhood effects at the municipal level in Denmark.

A final and substantial contribution of this thesis is to outline what can actually be learned about equality of opportunity from studies in intergenerational mobility and neighborhood effects. The majority of research in both intergenerational mobility and neighborhood effects concerns itself with equality of opportunity. With few exceptions (e.g. Roemer 2000 and

Roemer (2004)), they do not explain what their findings substantially tell us about extend of equality of opportunity. Drawing on three modern philosophical conceptions of equality of opportunity, I argue firstly that implicitly defining relations between equality of opportunity and mobility measures is insufficient to link the two. I secondly argue that most intergenerational mobility studies by nature are under-defined. They cannot tell us to what extent a society experiences inequality of opportunity. Some notions of equality of opportunity can be fulfilled for all levels of intergenerational mobility, while others are consistent with many different levels of opportunity. Additional information, such as differences in policies or contrasts between groups of individuals is necessary to link measured mobility to inequalities in opportunity. I thirdly argue that the existence of neighborhood effects on income mobility is a substantial deviation from relevant notions of equality of opportunity. This bears relevance for both scholars working on intergenerational mobility and neighborhood effects insofar as they are interested in determining levels and violations of equality of opportunity. It also relates the findings from investigations of the sub-questions to the overall research question of the thesis. The finding of substantial variation in mobility across Danish municipalities shows that there are general violations of equality of opportunity. The relations, however, do not appear to be causal at the municipal level, and so we cannot discern whether neighborhoods or individuals are driving the differences between municipalities in mobility and thereby equality of opportunity.

Equality of opportunity is at the center of our societies' foundations. In this thesis, I attempt to make contributions to our understanding of it. Limitations, of course, apply to the analysis. Among them are, firstly, that all measures of mobility, and neighborhood characteristics and effects are at the municipal level. Finding no clear indications of violations of equality of opportunity at this level does not imply that equality of opportunity is not limited by neighborhood characteristics at other levels of spatial aggregation. Secondly, all results hinges on the available data. In this thesis I make use of Danish register data made available through the IKE group at Aalborg University and provided by Statistics Denmark. The available data is limited on a per project basis. As a result, some variables of interest, such as local school quality and peer networks, have not been available. Thirdly, while the test for causal impacts of neighborhoods is rejected at the municipal level, this does not mean that none exist. The test is specified a priori, and may not cover all types of neighborhood effects at the municipal level. In future research, one can view the results from the test in an exploratory perspective to inform new causal effect tests.[2] The limitations notwithstanding, this thesis does make substantial contributions to the literature on intergenerational mobility and neighborhood effects.

1.1 A Roadmap

The thesis is divided into three main parts. In chapter 2, I in turn introduce three main conceptions of equality of opportunity in modern philosophical literature. In a final discussion I relate each notion to the concept of intergenerational mobility, and discuss the differences among them. This discussion leads to a basic framework of how differences in mobility within the country, and neighborhood effects in particular are indicative of inequality of opportunity.

Chapter 3 concerns the investigation of differences in mobility across Danish municipalities. In section 3.1 I introduce the main intergenerational mobility theory that has shaped the intergenerational mobility measurement literature, and chapter 3.2 contains discussions of the methodological considerations on measuring mobility that derives from it. Section 3.3 contains a description of the register data and choices I use to estimate intergenerational mobility at the municipality level, as well as the arguments for locating the analysis at the municipal level. The data introduced in this section is used throughout the empirical studies in the paper. Before estimating the variation of mobility at the municipal level, I provide national baseline estimates in section 3.4. I find that the baseline is consistent with existing research, and so continue to estimating the variation in mobility across different municipalities in section 3.5 where I also conclude on the findings on mobility at the municipal level, as well as their meaning in terms of equality of opportunity.

The purpose of chapter 4 is to explore the last two sub research questions: what can explain the differences in mobility across municipalities, what are the predictive capacity of municipal

[2]See e.g. Gelman and Loken 2014 and Tokey 1980 for this exploratory versus causal inference research approach.

level covariates, and is there evidence of causal effects of neighborhoods having an effect on intergenerational mobility, and in turn on equality of opportunity. Section 4.1 first contains an introduction to the neighborhood effects literature and the drivers of mobility at the neighborhood level that is inherent to it. It also provides an operationalization of the study. 4.2 gives an introduction to the studies of drivers of intergenerational mobility, including the studies that intersect the neighborhood effects literature and the intergenerational mobility literature. While some studies focus on neighborhood related factors, much prior research has emphasized factors other than the neighborhood, which suggests that the differences in mobility are in fact not driven by neighborhood characteristics, but perhaps self-selection into local areas. Section 4.3 describes the data used for the investigations of analysis of relations in section 4.4. The section shows that absolute mobility does in fact covary with many of municipal variables. I therefore turn to developing a causal inference framework in section 4.5, which I test empirically in 4.6. With the findings in section 4.6 I am also able to discuss what can be said about neighborhoods and equality of opportunity from this study. Chapter 5 includes a summary of the findings of the study, the limitations that apply, as well as derived questions for future research. Finally, chapter 6 concludes the thesis.

Chapter 2

Notions of Equality of Opportunity

In order to discuss equality of opportunity and intergenerational mobility and its measurement, it is helpful to understand what equality of opportunity is and what it implies. Such discussions are often not present in the economic literature on intergenerational mobility literature, even in review articles (these include Solon 1999, 2002; Blanden 2013; Torche 2015).[1] An example of the limited discussion of linkages between equality of opportunity and intergenerational mobility is found in the review article by (Black and Devereux, 2010, p. 3):

> Which brings us to the motivation for all of this research: What is the "optimal" amount of intergenerational mobility? Many people favour equality of opportunity as an underlying goal of society – the idea that poor children should have the same opportunities for success as rich children. Those who work hard should be able to succeed, regardless of family background. However, zero intergenerational correlation is not necessarily the optimum.

In the quote, Black and Devereux (2010) equate equality of opportunity with poor children having same opportunities as rich children. While this does resonate with a first approximation to the concept, this is a limited understanding of equality of opportunity, which does not help us understand what should be equalized, how, or why. A notable excemption to the

[1] Nor is it usually defined in appearances in political justifications (see for example Beck-Nilsson 2018; Regeringen 2007; UVM 2018), although this fact might be less surprising.

simplistic approach is the articles by Roemer (2000, 2004) who discusses a particular understanding of equality of opportunity and relates it to the intergenerational mobility literature. To the best of my knowledge, fundamental discussions of equality of opportunity are void from the neighborhood effects literature. This is so even if the literature is tackling questions that relate directly to inequality of opportunity, as I will discuss.

I draw on three modern notions of equality of opportunity from philosophy as surveyed by Arneson (2015, 2018) to provide a framework for discussing the notion of equality of opportunity.[2] In particular, I describe the notions of formal equality of opportunity and fair equality of opportunity, both of which are notable components of Rawls' 1999 theory of justice, and what Arneson (2018) terms the luck egalitarian notion of equality of opportunity to which Roemer (1998, 2000, 2004)[3] is a contributor.

As I discuss the notions of equality of opportunity of each position, I will make use of a hypothetical society to point out differences in each.[4] Let the hypothetical society be A. In A there exists three types of positions that confer rewards; ordered from highest to lowest rewards the positions are management, high value production, and finally low value production. In A lives multi-generation families. An individual's social (and economic) class follows from their position. Individuals in class a has management positions, individuals in b takes on high value production positions, and finally individuals in c take on low value production jobs. I further define intergenerational mobility as the relation between parental and child reward outcomes as a first approximation in this chapter. Under complete intergenerational mobility parental outcomes are statistically independent of child outcomes. Under complete intergenerational *im*mobility parental outcomes perfectly predict child outcomes. In addition, assume that neighborhood effects are characterized by being effects arising from qualities of the local neighborhood that either (1) changes a statistical relationship between parents and

[2]The idea of equality of opportunity has a long history in philosophic thought. It can, for example, be traced back to Plato's definition of an ideal state, Utopia. Here Plato requires that all children of his Utopia be given equal educational opportunity at the age of 10 (Durant, 2006, chapter 1, VII). It is noteworthy that Plato, according to Durant (2006), has an explicit purpose of utilizing all talent in the society. That is, the objective of equality of opportunity, for Plato, is to create an optimal state in the sense that people live harmoniously, having satisfied their individual psychological requirements.

[3]Roemer and his 1998 book *Equality of Opportunity* is, in addition, one of the few philosophical texts that some economists sporadically refers to (an example is a single footnote in Björklund and Jäntti 2012).

[4]This approach is inspired by the introduction in Solon (1999).

children's outcomes, or (2) changes child outcomes given parental outcomes. These distinctions allow me to discuss the relation of equality of opportunity to intergenerational mobility and neighborhood effects in the following sections.

2.1 Formal Equality of Opportunity

Formal equality of opportunity entails the following: For every position in society that confers a societal value, only characteristics relevant for the position can be considered when an individual is to be chosen for it (Arneson, 2015). Formal equality is also often referred to as non-discriminatory equality of opportunity (see e.g. Fleurbaey and Peragine (2013); Roemer (2002)), and appears as a minimum requirement for other notions of equality of opportunity (e.g. Rawls 1999, p. 57, 62). To exemplify the enactment of formal equality of opportunity in A, consider the following example of a high value production position. A firm needs a data scientist who can analyse data that the firm collects on its customers. Qualities relevant to the task include (1) statistical analysis skills, (2) knowledge of consumer theory, and (3) the ability to write reports. If formal equality of opportunity is fulfilled, any individual can apply for the position, and will be considered only on these three qualities. Formal equality of opportunity is not fulfilled if for example the ethnicity of the applicant is also considered by the employer in reviewing an application. Related examples include factors such as family relations (nepotism), gender, and religion.

The initial positions of an individual is irrelevant to formal equality of opportunity. A child of a rich business-owner who has gone to private schools and colleges is evaluated on the same criteria as a child who has grown up in a ghetto with little access to quality schooling or mentoring. It is clear that formal equality of opportunity does not imply any specific level of intergenerational mobility. Returning to A, assume that each job category (a, b, and c) has specific educational requirements. The cost of education for management positions is the highest, and the cost for low production positions is the lowest. In addition, assume that parents cannot afford education that will provide their child with the skills relevant for positions above the parents' position (but they can afford education relevant for their own position). If each set of parents attempt to maximize their child's future reward, then this society can have complete formal equality of opportunity, even if there is no intergenerational mobility - all children do what their parents do.

The fact that any level of intergenerational mobility is permissible under formal equality of opportunity suggests that this notion is not in line with the general idea described by Black and Devereux (2010) in the introductory quote. Black and Devereux are pointing to the fact that there appears to be a moral injustice in having children from poor and rich families fight battles on the same terms. But, the equal consideration for positions on merits is completely disregarding of what background the child comes from. In addition, the notion is also disregarding of the neighborhood a child grew up in, and how the child acquired skills relevant for a given position. What matters is whether or not the child has the skills.

2.2 Fair Equality of Opportunity

Fair equality of opportunity is a concept that stems from Rawls (1999), "*A Theory of Justice*".[5] Rawls' notion of equality of opportunity is embedded in his theory of justice. Unlike formal equality of opportunity, fair equality of opportunity is concerned not just with the competition for positions, but also with how each individual has been prepared for the competition.

To tackle Rawl's notion properly we need to understand the foundation for his theory of justice as the notion is embedded in it. Rawls develops his foundation from a hypothetical: Take all individuals and strip them of their characteristics (race, gender, wealth, nearsightedness, risk-aversion, etc.) so that they be in an initial state, where all are equal an no one knows their future position in society. In this initial state (behind what Rawls calls the 'veil of ignorance') the future members of society must decide on the preferred institutional of their society. Rawls finds that two main principles that must be considered in this order. The first is the principle of liberty, under which every person has the right to free speech, religion, and other basic freedoms. The second is the equality principle.[6] The latter principle consists

[5] The book was originally published in 1971 and has arguably made Rawls one of the most prominent political philosophy scholars of modern time, providing a political philosophical framework for considering what constitutes a just society in opposition to particularly utilitarianism (Sandel, 2010, 140-166).

[6] Rawls consider several specifications for his second principle. The one he finally applies is what he terms

of two main elements which must also be considered: fair equality of opportunity and the difference principle. The former concept is the one of primary interest here. (Rawls, 1999, p. 63) defines fair equality of opportunity as follows

> More specifically, assuming that there is a distribution of natural assets, those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial position in the social system.

Under this concept of equality of opportunity, no circumstances should affect the an individuals chance to participate in the societal competition for rewards. Only characteristics that are innate to the individual - such as inherent absorptive capacity, and willingness to exert effort - are allowed to affect the expected societal reward of the individual. Rawls' argues that by this principle, we have a just *procedure* for determining societal rewards, and as a result the outcomes of the procedure are morally just (Rawls, 1999, p. 73-78).[7]. Rawls' second equality principle, the difference principle, asserts that a societal system is only just if it is the one that maximizes the reward to the individuals least well off in society.[8] If two systems are similar in that both makes the least advantaged best of, then any configuration that makes others better off can be allowed (Pareto improvements are allowed). The important aspect of the difference principle for this chapter is that the difference principle acts as a filter on outcomes: Any outcome that follows directly from fair equality of opportunity can be altered for the benefit of the least worst off in society. Some individuals who are not in the bottom can even be made worse off for the benefit of the least advantaged (for example through taxation).

We now have a basic conception of Rawls theory of justice, and we can consider what his fair equality of opportunity would look like in A. In essence, fair equality of opportunity arises

democratic equality (Rawls, 1999, p. 57).

[7]Rawls relies on procedural justice to assert moral justification of outcomes. (Sandel, 2010, p. 159)) provides a critique of this position. By allowing for innate characteristics to affect an individuals expected outcomes, Sandel asserts that Rawls also allows for potentially morally arbitrary outcomes. An example is whether a child is first or second born. This may impact the child's willingness to assert effort and thereby expected outcomes under Rawls' Fair equality of opportunity. This, however, appears to be a lottery of birth and is therefore arbitrary.

[8]The principle is in a sense a maximin principle although (Rawls, 1999, p. 65-73) rejects the term due to its former use in optimization under uncertainty by von Neumann and Morgenstern (1944). The term has stuck, however, and has been used for example by Roemer (2002).

when circumstances such as family wealth or childhood neighborhood has no relation to child outcomes *given* innate abilities and talents, and inherent willingness to exert effort. Some inequality of expected outcomes can be present, and some of it can even be tied to the family. If, for example, we assume some genetic transmission of distribution of traits, say F_i (for $i = a, b, c$), are dissimilar across groups ($F_a \neq F_b, F_c$), then children from each group can have differing expected outcomes due to innate abilities. This can be the case even if there are no family investment in the child by parents in A or neighborhood characteristics (such as school financing or the amount of local crime) that differs. If differences in distributions of traits are correlated with parental expected positions and traits are inheritable, then positive statistical relations between parental and child expected outcomes can exist under fair equality of opportunity.[9]

The last paragraph in essence states that looking only at expected outcomes, different levels of intergenerational mobility from none to a large extend is possible. The level of intergenerational mobility will depend on the degree to which parental traits that are relevant for determining expected outcomes are inheritable. A similar case cannot be made for neighborhood effects. Neighborhood effects arise from characteristics of neighborhoods external to the child and family. From the quote above, it is clear that any characteristic not innate to the individual cannot determine expected outcomes under fair equality of opportunity. To exemplify this, assume that children in A live in different neighborhoods defined by their amount of local crime. Assume in addition that children's learning capacity is limited by anxiety arising from worrying about local crime. The children in A living with more crime as a result learn less than children living with less crime. If education matters for children's access to societal positions, then neighborhoods will affect expected outcomes. Such a case is not allowed if formal equality of opportunity applies.

Rawlsian fair equality of opportunity is concerned with expected outcomes. This inherently makes any operationalization of Rawlsian equality of opportunity challenging - we cannot determine what constitutes an expected outcome. In addition, what is measurable is the actual outcomes in society, which may differ from the expected outcomes. Adding to the problem of

[9] The difference should, however, disappear if we in some way could control for parental traits.

determining the extent of fair equality of opportunity of outcomes is the difference principle. Rawls' second equality principle ensures that outcomes may differ from what directly obtain under fair equality of opportunity. Take an example of A in which it has been decided that children with particular inherent traits should receive additional education for the benefit of the least advantaged (the government in A expects to be able to tax realized rewards). The actual outcomes in A follow from first ensuring fair equality of opportunity, secondly from changing the expected outcome distribution through policy, and thirdly through the randomness of realization of outcomes. Measuring only realized outcomes, one must be able to control for the effect of the second difference principle policies on top to infer whether fair equality of opportunity actually obtains (assuming that moving from realized to expected outcome is possible). The problem may be less severe in the case of neighborhood effects, as differences in expected values prior to application of the difference principle should be equalized. As a result, differential policy based on neighborhoods should not be driving realized outcomes.

The Rawlsian fair equality of opportunity is significantly more elaborate than the formal equality of opportunity. On its surface it appears to encapsulate what Black and Devereux (2010) suggests in their call. However, due to the abstract nature of the notion (dealing in expected and not actual outcomes) as well as potential filtering through policy mechanisms, it appears that investigations of whether or to what extent fair equality of opportunity obtains is a problem.

2.3 Algorithmic Equality of Opportunity

The final position that I investigate is that set out in John Roemer's "Equality of Opportunity" (1998). Roemer has since developed his position in contributions to the intergenerational mobility (Roemer, 2000, 2004) and inequality literatures (see e.g. Björklund et al. 2012). In many ways, Roemer takes a position that is similar to Rawls' when discussing equality of opportunity, but unlike the abstract Rawlsian notion, Roemer provides a higher degree of formalization and simultaneous flexibility in that underlying notions of justice can be used to specify relevant settings and notions of morality. This flexibility is encapsulated in Roemer describing his notion as an equality of opportunity 'algorithm' (Roemer, 2002).

Roemer (2004) relies on five different elements in discussing equality of opportunity: Objective, u, circumstances, C, Types, t, Effort, e, and Policy Instruments, φ. An *objective* is some measure of well-being that the government (or a benevolent social planner[10]) has an interest in equalizing to an extent. *Circumstances* are individual characteristics for which society does not wish to hold the individuals accountable when rewards (or outcomes) are conferred.[11] *Types* are specific constellations of circumstances into which we can group individuals. The number of types will be contingent on the relevant circumstances. *Effort* is the actions individuals undertake for which society does want to hold them accountable. Roemer assumes that for each type exists a specific distribution of effort, F^t, such that an individual performing at the highest possible effort level among poor, inner city child[12] will show lower effort levels than a child born to resourceful parents growing up in a wealth suburban area with well functioning education. Within type effort-percentiles is Roemer's main emphasis.[13] In many cases, Roemer defines effort as the residual determinant of outcomes given types. Finally, *Instruments* are the specific public policies of the society that can change individual values of the objective. Together these five elements constitutes the foundation for considering Roemer's equality of opportunity.

Roemer's equality of opportunity implies that outcomes are independent of type and depend

[10]The use of the Social Planner is an often occurring tool in economics. The social planner is an abstraction used in a setting where particular choices must be made in order optimize some chosen objective (Bowles, 2004)[p. 58]).

[11](Roemer, 1998, §2) makes a distinction between accountability and responsibility, where accountability confers rewards or punishments, whereas responsibility may not: An individual can be held responsible for any action, but accountable only for those performed under free will - those that are not determined by circumstances. An example is that of The Third Wave (Jones, 1976), in which 200 high school students in 1967 were lead by social forces (which we may term as circumstances) to unwittingly act as fascists for a full week. Under Roemer's notion, we do not want to hold the students accountable under these circumstances, even if we wish to hold them responsible for certain actions. Naturally, these distinctions are drawn with fine lines and will be situationally contingent.

[12]Inner city poverty is a prevalent problem in the US, documented by among others Sampson (2012) and Wilson (1987), that lead to the large scale Moving to Opportunity Project which attempted to move families out of poverty ridden inner city areas through vouchers (Gennetian et al., 2011).

[13]To distinguish within from between type consider the two students in our hypothetical society. One grew up in a wealthy family and went to prep schools, the other in a poor inner city family with low quality schooling (few mentors, lacking resources, etc). Assume that both children are top 5 in their class. This is their *within* distribution position. The between comparison of their positions would compare their positions in the total population of students (e.g. on test scores or type of college attended). It is likely that the inner city student fares much worse in this comparison due to Roemer's circumstances.

only on (within group percentile) effort, given the the policy instrument ($u(e, t, \varphi) = u(e, \varphi)$). This definition implies that Roemer explicitly can extract optimal policies (φ^{EOp}) when relevant objectives, circumstances and effort has been defined (and some strong assumptions, including monotonicity of effort and outcomes made). Betts and Roemer (1998) provide an example: using the National Longitudinal Study of Young Men, they investigate what level of educational spending per pupil (at age 16) would be needed to provide equality of opportunity for black and white males. Equality of opportunity is understood as equal outcomes for similar percentile effort levels in each type (black or white). Using among other things quantile regressions they find that the necessary spending must be 10:1 in favour of Blacks. This example (cited by Roemer 2000, 2002) also highlights a final important aspect of Roemer's approach: when deciding on relevant circumstances, a normative decision must be made as to what constitutes characteristics that we do not wish to hold an individual accountable for. Betts and Roemer chooses race as the only circumstance, but we can easily imagine more that we want to include. Roemer (2004) links the discussion of equality of opportunity to intergenerational mobility by defining four sets of sub-circumstances that we may wish to equalize for.[14] In order of likelihood that people may wish to equalize for it, these are (C1) Social connections, (C2) formation of beliefs and skills through culture and investment in the child, (C3) genetic transmission of ability,[15] and (C4) formation of preferences and aspirations by family influence. Roemer believes most people would agree that society must attempt to equalize for (C1) and (C2) in competition for rewards, whereas (C3) and (C4) are more contentious.

From the discussion of Roemer's equality of opportunity conception, it is immediately clear that neighborhood effects in most cases would fall into the category of circumstances. What is also clear is that the extent to which equality of opportunity is realized will depend on how the researcher (or social planner) determines what constitutes the set of complete rele-

[14]Whereas Roemer (1998) formerly has argued that one must account for differing abilities when equalizing, he here takes a more relaxed view that is closer aligned to that of Rawls.

[15]The extent to which socially and economically relevant abilities are inherited is questionable. Bowles and Gintis (2011), originally published in 1976, early on showed evidence of low importance of any inherited abilities in shaping educational outcome. More recent research in psychology, neuroscience, and economics has questioned to what extent ability is genetically inheritable, emphasizing that a nature-nurture dichotomy is insufficient as even genetic characteristics can be altered through external circumstances (Heckman, 2008a, p. 307).

vant circumstances. Nevertheless, determining whether neighborhood effects are a driver of outcomes is possible within the Roemerian algorithmic approach.

Consider now the application of Roemer's (later) notion of equality of opportunity in society A with the purpose of investigating intergenerational mobility: assume that the relevant outcomes are jobs (and their rewards), circumstances are at an aggregate level parental job category, and the policy instrument is public educational resources. Assume in addition, that sub-circumstances include that each child has some social connections through their parents, that parents have the opportunity to invest culturally and monetarily in their children, that children have some ability transmitted genetically from their parents, and finally that the parents can influence the child's preferences and aspirations. Under complete equalization, effort will depend on some innate characteristics of the child (potentially different by type), which is independent of parents. In this case, parent and child outcomes will be independent. If we assume that people in A decide that they only want to equalize for connections and deliberate investments in children, then sub-circumstances (C3) and (C4) will remain. If these increase the likelihood of ending up in the same outcome group as one's parent, then equality of opportunity allows for some connection between parental and child outcomes. This underscores the fact that Roemer's algorithmic approach is flexible, and in addition it questions what should be expected when measuring intergenerational mobility - should we want complete mobility (no relation), or allow for some liberty for parents to affect children? This is a normative question.

2.4 Discussion

The three notions of equality of opportunity differ in several ways, including their coverage of outcomes, underlying moral justifications, and most notably their applicability in empirical investigations. Formal equality of opportunity is the least relevant of the three positions as any levels of intergenerational mobility and neighborhood effects are permissible under this notion. The only requirement is that candidates for positions be considered *only* on their merits relevant to the position. Naturally, substantial research in discrimination (see e.g.

O'Flaherty 2015) has documented that even this, albeit simple, characteristic can be fulfilled to varying extends. Rawlsian equality of opportunity is stricter. Any characteristic that is not innate to the individual is not allowed to affect *expected* outcomes of any individual in the society. The question remains: where do we draw the line for what are truly innate, and what is in effect determined by circumstances. The result is that it can be difficult to determine the degree of equality of opportunity in the Rawlsian sense. Adding to the problem is the fact that Rawls is concerned with abstract expected outcomes, and not observed ones. As a result of the existence of the difference principle, we cannot be certain as to what level of statistical relationship between children and parents actually obtain a priori, unless the complete institutional schematics of the society is given. Similarly we also cannot directly back out the extent of equality of opportunity from a given level of intergenerational mobility. What is, on the other hand possible, is to determine that if neighborhoods somehow have a causal effect on observed children's outcomes (given parental background), or on the relation between children and parents outcomes, this must be a violation of the Rawlsian equality of opportunity, given that we know whether particular policies put in place to redistribute expected outcomes in the society, as well as the translating function from expected to actual outcomes. This is naturally not a given.

Roemer‘s algorithmic approach is more flexible - it is up to society (or researcher?) to decide what outcome is of interest, what circumstances are relevant, and what policy must be used to equalize across effort-levels within types. The flexibility of Roemer's approach is helpful for the researcher in intergenerational mobility in the following sense: Given all of the relevant information, such as the policy regime, and the relevant types, we can actually determine if variation in intergenerational mobility across areas violate intergenerational mobility. However, a problem arises as we must consider whether any *observed* spatial variation in mobility in effect does make for a violation. Take again A, and assume that workers in categories a, b, and c live in different neighborhoods. They pass on innate characteristics to their children (which lies in a spectrum that we will accept, including family culture and genetic capacity for effort). Each set of characteristics are translated into mobility different, so that each group has a distinct distribution of mobility ($M_i, i = a, b, c$) conditional on other circumstances. In this case, each local area will be characterized by different levels of intergenerational mobility as well. This is in essence the self-selection problem that is a returning feature in this study. Without knowing the likelighood of families self-selecting into areas, and what their innate characteristics are, we cannot determine from observed variation in mobility that there are violations of the algorithmic equality of opportunity. As for Rawls, however, a difference of case is made for neighborhood effects. Any effect that is entirely external to the family and individual, but which affects the child's outcome or the relation to parental outcome, *can* be a detectable violation of equality of opportunity. Unlike the Rawlsian position, which categorically rejects neighborhood effects, the Roemerian position, however, needs a moral justification for terming a neighborhood effect a circumstance. If we are willing to provide this, then we have a case for arguing that neighborhood effects are violating equality of opportunity.

From this discussion arises two main elements. Firstly, notions of equality of opportunity are sufficiently distinct that when we want to discuss equality of opportunity, and in particular measure the extent to which it obtains in a given society, we must be explicit about what we understand by the concept. Secondly, even limiting ourselves to a subset of notions that seems relevant from a modern societal perspective (the Rawlsian and Roemerian approaches), both require substantial knowledge about underlying differences in innate characteristics, and implemented policies to determine whether or not equality of opportunity obtains. Any simple intergenerational mobility measure is by nature under-defined in providing information about equality of opportunity. Thirdly, with the Rawlsian and Roemerian conceptions of equality of opportunity, we can in fact not argue that a violation of equality of opportunity exists even if there is variation in intergenerational mobility across area. This comes about primarily from the fact that both positions allow for some variation in expected and actual outcomes (for Rawls and Roemer respectively) due to innate characteristics. If the distribution of innate characteristics is inheritable, and correlates with spatial distributions of families, then differences in observed mobility can be permissible. It is, however, the case that we can argue from a Roemerian perspective that any observed causal neighborhood effect actually is a violation of intergenerational mobility. In the Rawlsian perspective, this would in an abstract perspective also be true (fundamentally, Rawls would not allow for differences in expected outcome due to neighborhoods), but we cannot simply move from observed outcomes to discussing intergenerational mobility.

The following propositions appears: We cannot directly use the under-defined equality of opportunity measures that usually apply to discern violations of equality of opportunity. We can, however, if we take a Roemerian approach, and specify neighborhoods as circumstances, show that any causal effect of neighborhoods in effect is a violation of the equality of opportunity that we are interested in. This is the foundation that I start from in this study. Before turning in the next chapters to the discussion of whether or not there are indications of violations of equality of opportunity, a relevant question that I have not addressed in this chapter is what constitutes relevant outcomes. Is income the relevant measure? Should we rather be concerned with fundamental freedoms (Sen, 2006), or should we, in a Kantian perspective, consider willingness to act morally as a better measure of a good life (Sandel, 2010, 117-123)? A deeper discussion of these notions require a further dive into philosophical notions of what constitutes good societies and values, a dive which is not possible here. In the economics discipline, income has received particular attention as the relevant outcome (or objective) when discussing economic mobility. In the following sections I follow this literature in focusing primarily on income, guided in large part by the seminal article by Becker and Tomes (1979) which emphasizes income, in line with former economic research, as it is assumed to transform directly to utility through consumption. In this thesis I follow this line of research in emphasis mobility in income terms, even when discussions of other relevant factors for a good life can be relevant.

Chapter 3

Intergenerational Mobility and Its Spatial Variation

This chapter has the purpose of investigating the observed spatial variation of intergenerational mobility in Denmark. In the first section, I introduce what has become the standard theory of intergenerational mobility, which I will draw upon to provide a basic framework for understanding within family mobility and what we can actually measure. The second part takes on the primary work in the literature, which concerns itself with the estimation of mobility. I describe a variety of measurements before implementing them first at the national level, and then at the municipal level. At the municipal level, I document substantial spatial variation in intergenerational mobility.

3.1 Standard theory of intergenerational mobility

Most of the literature on intergenerational mobility has been empirically oriented. In this section I introduce the main theoretical framework that have been used as a reference point for much of this empirical work in 1990's, and the 2000's. The purpose of introducing the model is to motivate the reason why, in the first place, we might expect that there be any relation between parents and childrens outcomes measured as income. It is noteworthy that more recent contributions to the literature, such as Chetty et al. (2014); Landersø and Heckman (2017) have taken a broader theoretical view, drawing on sociology (e.g. Wilson (1987)) and broader socio-economic research (Landersø and Heckman (2017) explicitly draw on Currie (2001) and Heckman et al. (2013)). Nevertheless, the human capital investment models are sufficiently constitutive in the literature that they merit explanation.[1]

The original economic model attempting to explain intergenerational mobility was developed by Becker and Tomes (1979), and later extended by Becker and Tomes (1986). The basic concept of the model is that parents find utility in both their own and children's (consumption) outcomes and can choose to invest in their children in order to raise their future economic returns. The income available to each individual depends on their human capital resources which (translated through a production function) produces income in the market economy. The human capital resources can either be provided through government investment, or through parental investment. The parents will choose how much to invest in their child based on perceived marginal return on the investment (i.e. how well will the child be at utilizing the human capital investment in the market), a characteristic inherent to the child. Under assumption of perfect credit markets (full information and no constraints on credit) parents can even borrow in their children's future income to make optimal investments in them. Solon (2004) provides an updated version of the model which explicitly considered government investment and linearity, relating it to the extensive literature using intergenerational elasticities to estimate intergenerational persistence, and Solon (2014) another update building on Solon (2004) by including transmissions separately from grandparents to children. Mogstad (2017) notes that this is the most recent substantial development of the intergenerational mobility model. Other recent modelling endeavours have put stronger interest in the early childhood line of research (see Francesconi and Heckman 2016) and investigate parental e.g. time spend with the children (Del Boca et al., 2014) on child development in classical optimization frameworks with varying productivity of investments over age-curve and cumulative investment. These models, are less directed at the intergenerational mobility literature

[1]The human capital literature took off from Becker (1993), originally published in 1964. Human capital, as explained by Becker, is a concept that captures marketable characteristics that increases productivity, and so can earn returns in a market economy. An alternate expression of human capital is that expressed by Bowles and Gintis (2011) in which we can view human capital as the educationally instilled capacity to obey orders in a hierarchical workspace (see Acemoglu and Autor (2009), chapter 1).

and more directly at the literature on child development.[2] I will use the model presented in Solon (2004) as an exemplary case of the class of intergenerational mobility models. In chapter 4 I will use a version of this model as a starting point for the consideration of causal effects.

The main features of Solon's (2004) model is that the relation between parents and children's outcomes increases in heritable characteristics that increases income, decreases in progressivism of public investment in children's earnings capacity, and finally it increases with the returns to investment in earnings capacity. The model follows the usual structure of developing a choice situation, which can be solved through choices made to optimize a certain objective (which is usually the abstract utility). To develop the Solon-model, assume that in a society, A, each family i has a parent generation denoted $t-1$, and a child generation denoted t. Within each family, the choice is made by the parents who must allocate their (after tax (τ)) income, $(1-\tau)y_{i,t-1}$ (they are not allowed to borrow in this model). It is assumed that the taxation is linear (that is, non-progressive). As stated for the Becker-Tomes model, each family's parents must allocate their income between the parent and child generations through present consumption $C_{i,t-1}$ and children's future consumptions $C_{i,t}$. The parents, however, cannot directly spend their income on the child's consumption (in adulthood), and therefore makes human capital investments $I_{\text{i,t-1}}$. The total budget constraint is then (assuming that parents spend all of their lifetime permanent income) of the parents is

$$(1-\tau)y_{\text{i,t-1}} = C_{\text{i,t-1}} + I_{\text{i,t-1}} \tag{3.1}$$

All taxes goes to the government that can redistribute these through spending, $G_{\text{i,t-1}}$. As this in essence is a human capital model, the spending naturally goes to human capital investment (we can think of provision of resources for kindergarten care, primary education, and free provision of higher education, in addition to health care and other services, for the Danish case).

With government expenditure and parental investment we can specify how children's human

[2] While not directly specified for the intergenerational mobility literature, this will be possible by extending the models from end of childhood into adolescence and adulthood, as noted by Del Boca et al. (2014).

capital is formed - Solon (and Becker and Tomes 1979, 1986) assumes that human capital depends positively with diminishing returns to the the sum of government spending and parental investment, as well as on some transmission of endowed attributes, $e_{\text{i,t-1}}$,

$$h_{\text{i,t-1}} = \theta \log(I_{\text{i,t-1}} + G_{\text{i,t-1}}) + e_{\text{i,t-1}}. \tag{3.2}$$

In equation 3.2 θ is the marginal return on investment in human capital. The $\log(\cdot)$ function ensures positive and diminishing marginal returns.[3] For attributes, Solon refers back to (Becker and Tomes, 1979, p . 1158), who state that these "*determined by the reputation and 'connections' of their families, the contributions to the ability, race, and other characteristics of children from the genetic constitutions of their families, and the learning, skills, goals, and other 'family commodities' acquired through belonging to a particular culture. Obviously, endowments depend on many characteristics of parents, grandparents, and other family members and may also be culturally influenced by other families.*" This description of $e_{\text{i,t-1}}$ further suggests that the term should depend on former periods values. Solon, in line with Becker and Tomes (1979) specifies a simple first-order autoregressive function for $e_{\text{i,t-1}}$ in order to catch the dependence:

$$e_{\text{i,t-1}} = \delta + \lambda e_{\text{i,t-1}} + v_{\text{i,t-1}}. \tag{3.3}$$

The term $\delta \in [0, 1]$ shows the rate of transmission of heritable attributes, while $v_{\text{i,t-1}}$ adds white noise to the process. Naturally, this term includes almost all of the elements I have discussed in relation to Rawls' and Roemer's understanding of equality of opportunity. I will return to a discussion of them later on.

The child's earnings are given by a simple production function depending only on their human capital, where p indicate returns to human capital, and μ the indicated minimum earnings.

$$\log y_{\text{i,t-1}} = \mu + p h_{\text{i,t-1}} \tag{3.4}$$

Parents' choice depends on the utility they derive from their own consumption, and the income

[3] that is, $\frac{\partial h_{\text{i,t-1}}}{\partial I_{\text{i,t-1}}} > 0$, $\frac{\partial^2 h_{\text{i,t-1}}}{\partial I_{\text{i,t-1}}} < 0$, and similarly so for $G_{\text{i,t-1}}$.

of their children. Solon chooses a standard Cobb-Douglas representation of the utility, with an 'altruism' parameter $\alpha \in (0,1)$ indicating parents preference for their children's consumption over their own (and also ensuring constant returns to scale, and diminishing marginal utility on both inputs)

$$U_i = C_{\text{i,t-1}}^{1-\alpha} y_{\text{i,t-1}}^{\alpha}. \tag{3.5}$$

By taking the logarithm on both sides, the expression becomes

$$u_i = (1-\alpha)\log C_{\text{i,t-1}} + \alpha \log y_{\text{i,t-1}} \tag{3.6}$$

Under the assumption that parents are aware of how the knowledge production takes place, as well as the parameters included in the process, then we can substitute in equations for $y_{\text{i,t-1}}$, and for $C_{\text{i,t-1}}$ by solving for it in 3.1. From this, the utility function can be simplified to a single choice variable, $I_{\text{i,t-1}}$ (remember that parents formerly had to choose both their own level of consumption and their investment in the child)

$$U_i = (1-\alpha)\log[(1-\tau)y_{\text{i,t-1}} - I_{\text{i,t-1}}] + \alpha[\mu + \theta \log(I_{\text{i,t-1}} + G_{\text{i,t-1}}) + pe_{\text{i,t-1}}] \tag{3.7}$$

To find the optimal level of investment in the child (seen from the perspective of the parents) we find the first order condition,

$$\frac{\partial U_i}{\partial I_{\text{i,t-1}}} = -\frac{(1-\alpha)}{[(1-\tau)y_{\text{i,t-1}} - I_{\text{i,t-1}}]} + \frac{\alpha\theta p}{I_{\text{i,t-1}} + G_{\text{i,t-1}}} = 0, \tag{3.8}$$

from which we can solve for

$$I_{\text{i,t-1}} = \left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}} - \left[\frac{1-\alpha}{1-\alpha(1-\theta p)}\right]G_{\text{i,t-1}}. \tag{3.9}$$

This is the choice parameter of the model (using it, we can recover the level of parental consumption as well). We see that higher parental income increases the amount of investment in the child, whereas government resources acts a part substitute. The extend of the crowd-out will depend both on parents altruism towards their children (α), the productivity of human capital investment (θ), as well as the return on human capital (p).

Finally, with the parental investment (and consumption) behaviour pinned down, we can define the relationship between parents and children incomes in the model. Substitute in to include the dependence on human capital, and we get

$$\log y_{\text{i,t-1}} = \mu + p[\theta \log(I_{\text{i,t-1}} + G_{\text{i,t-1}}) + e_{\text{i,t-1}}] \tag{3.10}$$

We can then plug in the optimized investment behaviour equation for the parents and derive the following expression (see appendix A.3):

$$\log y_{\text{i,t-1}} = \mu + \underbrace{p\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)\right)}_{A} + p\log\left(y_{\text{i,t-1}}\right) + \underbrace{p\log\left(1+\frac{G_{\text{i,t-1}}}{(1-\tau)y_{\text{i,t-1}}}\right)}_{B} + pe_{\text{i,t-1}}. \tag{3.11}$$

The resulting equation contains five parts: μ, the term A, which in essence is a constant determined by parameters in the model, the term $p\log y_{\text{i,t-1}}$ which gives us the first approximation to a coefficient in a regression, and a term, B, that depends on the relative share of government expenditures to private income. Finally the term $pe_{\text{i,t-1}}$ includes inherited characteristics. Next, Solon (2004) asserts that we should have a look at B. By letting $\frac{G_{\text{i,t-1}}}{(1-\tau)y_{\text{i,t-1}}} = r$, we can imagine the share of government resources per family i to parental after-tax resources as an percentage r. We can show by a Maclaurin Expansion that for small shares of government to total family income, B can be reasonably approximated by r.[4] Solon further rewrites B by linearising the term: $B \approxeq \varphi - \gamma \log y_{\text{i,t-1}}$. For $0 < \gamma$, where γ is an expression of the elasticity of government expenditures to income, the approximation assumes that as after-tax income increases, the government share of expenditures decreases as a share of total investment in children in the family. φ sets a baseline level of expenditures per family. We

[4] The MacLaurin expansion is a polynomial approximation of a function $f(x)$, expanded around the point $x = 0$ (see Chiang and Wainwright 2005, p. 252). The expansion takes the form $f(x) = \frac{f(0)}{0!} + f(x) = \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^n(0)}{n!}x^n$. Letting the function $f(x) = \log(1+x)$, we can find an approximation around 0: $f(x) \approx 0 + x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \ldots$. For small x (< 1 and close to 0), $\log(1+x)$ is reasonably approximated by x. as higher order terms < 1 multiplied by fractions less than one increases to zero fast, and in addition partly cancel each other out.

can now express the relation between parental and child incomes as

$$\begin{aligned}
\log y_{\text{i,t-1}} &= \mu + p\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)\right) + p\log(y_{\text{i,t-1}}) + p\varphi - p\gamma \log y_{\text{i,t-1}} + pe_{\text{i,t-1}} \\
&= \underbrace{\left[\mu + p\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)\right) + p\varphi\right]}_{\mu^*} + p(1-\gamma)\log(y_{\text{i,t-1}}) + pe_{\text{i,t-1}} \\
&= \mu^* + p(1-\gamma)\log(y_{\text{i,t-1}}) + pe_{\text{i,t-1}}
\end{aligned} \tag{3.12}$$

This final equation approximate what has often been estimated in the intergenerational mobility literature. However, as Solon further notes, the last term here is not an *iid* error-term, put the inherited characteristics of the child,

$$e_{\text{i,t-1}} = \delta + \lambda e_{\text{i,t-1}} + v_{\text{i,t-1}}. \tag{3.13}$$

By noting this, we also see that as child's income depends on parental income, so must parental income depend on parental endowments. This means that there will be a correlation between $\log y_{\text{i,t-1}}$ and $e_{\text{i,t-1}}$. As a result, any regression of log of child income on parental income of the form $\log y_{\text{i,t-1}} = \alpha + \beta \log y_{\text{i,t-1}} + \epsilon$ would, following the above equation, also attribute to the regression coefficient β all the influences that follow from serially correlated characteristics within the family. Solon finds a stationary process coefficient from the regression of the following form:[5]

$$\beta = \frac{(1-\gamma)\theta p + \lambda}{1 + (1-\gamma)\theta p\lambda} \tag{3.14}$$

For reasonable values of each parameter, we find that β is increasing in the returns to education (P), the productivity of human capital production (θ), and the extend to which characteristics are inherited (λ). Finally, it is decreasing in the speed at which personal income crowds out the importance of government spending (γ).

This is the basic structure of the model outlined by Solon (2004). It is interesting to note that

[5] See Greene 2011, p. 919 for the derivation.

by letting some of the parameters in the model vary by parental income level, we can create non-linearities in the model. For example, let us assume that heritability is not constant across income levels, but that the importance and likelihood of using parental connections grows with income. In that case, β will be increasing over parental income (see an example in figure A.2 in appendix A.2).

We can relate this model to Rawls' and Roemer's notions of equality of opportunity. It is clear that in this canonical model, the outcome of interest is income. The alternate measures of socio-economic status considered in the sociological literature, such as occupation and education (Torche, 2015), differ. For the permanent income measure, a consumption-utilitarian understanding of well-being, mixed with altruism is introduced. Parents care for their kids, and are willing to forego parts of their own immediate consumption driven well-being in order to provide for better outcomes of their kids.[6] The investments available to each child, however, follows directly from parental income. Except, the government can attempt to prevent part of this transmission of parental income through investment into the child by providing a additional resources. This follows in line with Roemer's notion of a policy instrument (however weak it may be stated in this theory). The heritable traits included in Roemer's and Rawls' discussions also appear in the model, all of which will influence the β. Naturally, if β is zero in this model, this means either that heritability and parental investment both are irrelevant for child outcomes,[7] or that public investment in children is sufficiently strong to crowd out parental investment *and* make up for heritable traits. In all likelihood, we should not expect $\beta = 0$ for any society, and both in Roemer and Rawls notions of equality of opportunity, particular elements that could be described as heritable (counted in willingness to assert effort) may be quite morally justifiable.

It is noteworthy that Willis (1986) asserts that Becker and Tomes (1986), one of the seminal theoretical pieces in the economic intergenerational mobility literature, actually understands

[6] Recent findings by experimental economists (often in collaboration with other social scientists) has emphasized that *homo economicus* can have strong altruistic behaviours (see e.g. Fehr and Schmidt 2006 for an early overview and potential explanations). The approach shown here follows in a neoclassical line in that it directly assumes children receives utility from parents may loose utility from lowering their own consumption, they balance this loss against utility gained from investing in their children (Simon, 1992).

[7] We see that if human capital investment by parents (or the government) had no effect, $\theta = 0$, then $\beta = \lambda$. If heritable trades had no effect, then $\beta = (1 - \gamma)\theta p$.

equality of opportunity as one somewhat similar to that of Rawls' complete second principle: namely that equality of opportunity is a situation *"an individual in a given generation would be indifferent with respect to the family into which he is born"* (Willis, 1986, p. 41). It may even be stronger yet in that it can requires no statistical relation at all between parental and child outcomes in our hypothetical society. The only differences that remains between the outcomes of individuals would be due to stochasticity, i.e. luck. Willis himself, on the other hand, appears to be confusing what defines an individual's 'own' resources and innate personal and social endowments when stating his version of equality of opportunity (one that he ascribes to Becker's Woytinsky lecture). Namely, under perfect markets (especially financial markets) any individual can borrow against human capital endowments to invest in one self and thereby realize his or her potential. If forces

A problem then arise: What is 'the right' amount of equality of opportunity? The question, rightfully, often appear in introductions to articles that discuss measures or attempt to measure equality of opportunity, (see e.g. Corak 2013 and Solon 1999). It is immediately obvious, that we cannot absolutely discern what would be an optimal β, firstly as it depends on several parameters; we can have several specifications, for example with abhorrent inequality due to extreme returns to education coupled with extreme government intervention, or extreme heritability of positions through social connections, but also with high government investment rates, or even a totalitarian state determining individual outcomes completely. All of these outcomes create low β values, but all may on ethical grounds seem unpalatable according to modern understandings of just societies where more or less (dependent on institutions and ideological preferences) government intervention generally is accepted in order to equalize for opportunity Alesina et al. (2017).[8]

The general approach that has been taken to this question of 'what is the right β' is a pragmatic one: We must start by gathering evidence to understand what constitutes the actual levels of mobility, and then investigate what society constructions can explain the found

[8]Naturally, different theories of justice and more extreme ideological orderings than accrue to the majority of western society's members can support these intervention schemes. One example is the communist society called for by Karl Marx in his Manifesto of the Communist Party. Several main themes of the imagined classless society where individual capital ownership is abolished would suggest a theory of justice that supports extreme government intervention, partly through heavy progressive taxation, and through free education for all children in public schools (Marx and Engels, 1978, p. 490).

estimates. (Solon, 1999, p. 1762), describing two societies with similar income inequality, one having complete mobility and the other no mobility, writes:

> *Although societies A and B have the same measured inequality within a generation, the two societies are tremendously different in the character of their inequality. And, once that is agreed to, we should wonder where our own society's intergenerational mobility lies along the spectrum between societies A and B, and how it compares to the mobility in other societies. Beyond that, we would like to know why each society has the degree of mobility it does. Even with all that information, reasonable people still would disagree about the fairness of their society's degree of mobility and about what, if anything, should be done about it. But without such information, it is difficult to have an informed opinion.*

In this paper I follow in this stream of literature and primarily take on investigations of what characterizes the measured intergenerational mobility. Naturally, as I have argued in chapter 2 and in this section, there are fundamental problems with respect to asserting that any specific levels of measured transmission relates directly to any other specific level of equality of opportunity. A starting point, however, is to follow the prescription, to investigate whether there in the first place are any indications of causal effects of neighborhoods. If there is, then we can proceed with the analysis, and attempt to investigate whether specific characteristics in line with the neighborhood effects literature have any meaningful statistical relationship to the data, and we can approach the question of whether neighborhoods in fact do matter for equality of opportunity.

3.2 Measuring Intergenerational Mobility

In this section I introduce several measures of intergenerational mobility that has been used in the literature related to the former section. While the intergenerational elasticity (the coefficient from regression of log child income on log parent income) is the measure that directly translates from the model framework I present, several of the related measures, including the rank-rank correlation and absolute mobility measured by transition matrices, can be understood from the model in the last chapter. Namely, parental investment in the child, heritable traits, returns to education, as well as public policy can all impact measured mobility. In particular, all measured can be used to investigate of mobility varies with local areas. I put substantial emphasis on the problems of measurements, as incorrect measurement can attenuate real differences in mobility so that we might not detect them even when they exist.

Intergenerational Elasticity

Mobility has been investigated with a range of different statistics. The most frequently used in former economic studies is the Intergenerational Income Elasticity (IGE) (Black and Devereux, 2010; Blanden, 2013; Corak, 2006b; Solon, 1999, 2002; Torche, 2015), the coefficient obtained from regressing log of child income, $\log y_{\text{i,t-1}}$ on log of parental income $\log y_{\text{i,t-1}}$. Let Y_c and Y_p be the logarithm of child and parental income, and we can then write the regression model as

$$Y_c = \beta_0 + \beta_{IGE} Y_c + \epsilon. \tag{3.15}$$

The regression provides a relative mobility estimates, the expected percentage increase in children's earnings from a percentage increase in parents income. A higher IGE suggests a stronger relation between children's income and their parents' income levels. The regression relates to the Solon (2004) model, and social mobility is generally assumed to be inversely related to the coefficient: the higher the IGE, the lower equality of opportunity.[9] Naturally, estimating equation 3.15 does not give the causal impact of parental income on child income. As indicated by our model, many variables are part of the error term, ϵ. To the extend that these variables correlate with parental income ($cor(Y_c, \epsilon) \neq 0$), the estimate of β_{IGE} will be incorporating omitted variables bias. Equation 3.14 shows that these may include heritability of family characteristics from parents (including culture, genetics, networks, and more), in addition to governmental impacts. As a result, direct interpretation of this measure as a pure measure of relation between incomes is misguided. Nevertheless, the estimate has been

[9]Corak (2006b) provide an alternate understanding of the IGE. By taking the anti-log of the estimated equation, one obtains what Corak describes as a measure of economic advantage. The higher the IGE, the higher the level of economic advantage, understood as the increase in expected income in the next generation for the family compared to families at lower levels of income.

extensively used in the literature, and as a general measure of equality of opportunity. In the following discussion I assume that what we actually want to measure is this general measure, and not the causal impact of earnings.

When estimating IGE's, the general assumption has been that the measured income should be *permanent income*, a measure of life-time averaged income (Solon, 1992). The concept of permanent income dates as far back as Friedman (1957) and corresponds to the original neoclassical foundations of the Solon (2004) models.[10] For the discussion that follows, I will assume that it is indeed permanent income and the proxied permanent economic status that is of interest. Several problems have been noted in estimation of IGEs; Measurement error due to transitory income shocks, life cycle bias in income, bias due to varying income dispersion, and non-linearities.[11]

Individual income may vary over the life time due to temporary shocks (due e.g. to short term unemployment, bonuses, illness, etc). Solon (1989, 1992) introduced measurement error from these shocks to the IG literature. Under the assumption that the shocks are transitory, we can view parents' and children's income true log income, Y_c^* and Y_p^* as measured with

[10] It is, more or less explicitly, assumed that both economic status and economic behavior should follow from permanent income. It is, however, also not clear whether either of these notions are sufficient for arguing for the use of permanent income interpreted as a steady life time income perspective. Evidence from behavioural economics has suggested that people in generally do not plan based on their expected life-long income, but rather that they act in accordance with mental accounting strategies, and exhibit bounded rationality (for an overview of the developments of the field of behaviour economics, see Thaler (2016)). In the Solon model we assume that parents invest in their children based on their income. If they invest in accordance with their perceived income, and the attached mental accounts, then we can assume that the investment behaviour to some extend will depend on present resources. Naturally, to the extend that borrowing is possible, parents may still do so, but this suggests that income when the child is young (and perhaps wealth) is a stronger determinant of potential life-time income. In addition, much of the research on early childhood development (see Heckman et al. (2013)) suggests that the environment that children grow up in, including their exposure to high quality early child care, has large effects on their future outcomes. If this is the case, and what we are interested in, is the transmission of parental characteristics to children, then we should be less interested in life-time economic status, proxied by permanent income, as opposed to parental status when the child is growing up. Together these two elements of critique suggests that we must make a choice when deciding whether to investigate transmissions between parents and children, or broader relations of permanent characteristics. If the former is of interest, then measuring parental characteristics (including income) over the period when the child is young may be a better measure. If the latter is of interest, then measures catching permanent characteristics are preferred.

[11] Another problem, which Solon (1989, 1992) shows can produce (downward) attenuation bias in IGE estimates is homogeneous sampling of children when full population data is not available. This, however, is not a problem in the present study, where I use administrative data covering the complete population.

identically, independently distributed random errors, $e_{\text{i,t-1}}$ and $e_{\text{i,t-1}}$ as

$$\begin{aligned} Y_c &= Y_c^* + e_{\text{i,t-1}} \\ Y_p &= Y_p^* + e_{\text{i,t-1}} \end{aligned} \tag{3.16}$$

Inserting into what we assume is the correct model, $Y_c^* = \beta_0 + \beta_{IGE} Y_p^* + \epsilon$, this gives

$$\begin{aligned} Y_c &= \beta_0 + \beta_{IGE} Y_p + \epsilon + e_{\text{i,t-1}} - \beta_{IGE} e_{\text{i,t-1}} \\ &= \beta_0 + \beta_{IGE} Y_p + \epsilon_1 - \beta_{IGE} e_{\text{i,t-1}} \\ &= \beta_0 + \beta_{IGE} Y_p + \nu \end{aligned} \tag{3.17}$$

Now, ν is a compound error-term. If the measurement error of children is truly independent, then this additional measurement error should pose no problem (but raise variance of the estimated coefficients). The parental measurement error, however, is directly related to the measured parental income. This implies that $cor(Y_p, \nu) \neq 0$ (we have already seen that this is the case, but we now add another layer), and the estimator will experience (downward) attenuation bias as the probability limit becomes

$$plim \hat{\beta}_{IGE} = \beta_{IGE} * \frac{\sigma^2_{Y_P}}{\sigma^2_{Y_P} + \sigma^2_{e_{\text{i,t-1}}}}. \tag{3.18}$$

As the variance of parental permanent income $\sigma^2_{Y_P}$, is less than the the sum of the variance of parental income, and parental measurement error $\sigma^2_{e_{\text{i,t-1}}}$, this is smaller than β_{IGE} with no classical measurement error. One way to deal with this type of attenuation bias is to average over several years of income (Solon, 1992). Another is to use IV methods to rid of the biased variance. As noted by both Black and Devereux (2010) and Solon (1992), however, the exclusion criterion[12] is not likely to be fulfilled for most conceived instruments in the literature, a common of which is parental education. Several authors have investigated the extend of attenuation bias due to transitory shocks, and Mazumder (2005) shows that for strong autocorrelation in the persistence of shocks (which adds additional complications), it can be necessary to use up to 20-30 years of observations to adequately average out parental

[12]In a simple regression, $y = \beta x + e$, where x is endogenous ($cor(x, e) \neq 0$), assume that z is an instrument. The exclusion restriction states that $cor(z, e) = 0$. This can be violated if z independently from x correlates with y.

income shocks. By using longer parental income estimates, he shows that US IGE's rise from about .4 to .6. Recent research (e.g. Landersø and Heckman 2017) have generally used less years, but nevertheless do concern themselves with measurement errors.

Life cycle bias is a related type of bias for the IGE estimator. The problem got particular attention with Grawe (2006) and Haider and Solon (2006), and has recently been surveyed by Nybom and Stuhler (2016).[13] Summarized shortly, life-cycle bias arise due to changing incomes over the life cycle - when individuals are young, they tend to have lower incomes. Around mid-life the income reaches it's peak, before in later life, the incomes decrease again. Added complications follow from heterogeneity; we may expect that individuals with shorter educations start working early and experience slow gradual increases of income over the life cycle, while individuals who attend university will experience later and likely steeper income increases. This problem will especially be pronounced in Denmark, where average ages of graduation from (short) vocational education and (long) university education is higher than in the US. A simple example of the life cycle bias under homogeneity follows from expressing children's log income at age t as $Y_{c,t} = \eta_t Y_c + v_t$, where η_t is the age t deviation form permanent income, and v_t is an iid error term. Inserting the original regression model into this expression, we get

$$Y_{c,t} = \eta_t \beta_{IGE} Y_p + (\eta_t \epsilon_i + v_t). \tag{3.19}$$

Naturally, of child income is underestimated ($\eta_t < 1$), this would suggest running the regression underestimates the true β_{IGE}. The problem naturally also extends to parental income, where Haider and Solon (2006) shows that under assumption of a model $Y_{p,t} = \eta_t Y_p + v_t$, the bias of the estimated IGE will take the form

$$plim \hat{\beta}_{IGE} = \beta \frac{\eta_t \sigma^2_{Y_p}}{\eta_t^2 \sigma^2_{Y_p} + \sigma^2_{v_t}}. \tag{3.20}$$

Depending on the extend of classical measurement error (from v_t) and life-cycle bias η_t, the actual estimated variance can be either over- or understated.

[13] Mazumder (2005) provides a related discussion of the importance of the age of measurement of parental income, showing that estimation bias related to parental age is smallest when using measurements of income around age 40.

Grawe (2006) and Haider and Solon (2006) suggest that one solution to overcoming life-cycle bias can be to estimate the parameters in the above function and adjust the estimated IGE estimate. This approach, however, is highly contingent of the specific parametrization, which also assumes homogeneity of bias. An alternate approach, suggested by Mazumder (2016) and Nybom and Stuhler (2016) is to use averaged measurements of income measured near mid-life. This approach would appear less error-prone as it does not make parametric assumptions on the life cycle bias. The question then becomes to investigate earnings when income profiles stabilize near permanent income levels. Nybom and Stuhler (2016) suggests that this happens around the mid-thirties in Norway, and Landersø and Heckman (2016b) show similar patterns for Denmark.[14]

IGE estimates may also vary due to varying dispersion of income. Any OLS regression coefficient can be decomposed into correlation and a ratio of standard deviations. Let a model be $Y = \beta X + e$, where β is estimated by OLS. Then, for a simple regression function, letting σ_Y and σ_X be standard errors of Y and X, σ_{XY} the covariance between the two variables, and $\rho = \sigma_{XY}/(\sigma_X \sigma_Y)$ their correlation, we can write

$$\beta = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \frac{\sigma_Y}{\sigma_X} = \rho \frac{\sigma_Y}{\sigma_X} \tag{3.21}$$

This implies that if the income distributions between parents and children vary, then the estimated IGE's will also vary (Black and Devereux, 2010). Landersø and Heckman (2017) investigates how Danish and US IGE's depend on spread in income, finding that large shares of Danish and US IGE estimates follow from intergenerational differences in income distributions. They generally find that children's earnings distributions show higher dispersion than parents, resulting in IGEs larger than log earnings correlations. They also show that choices of income, and their related distributions drive a big part of the differences between e.g. total income and earnings IGEs. I will return to the question of what income measures to use later. No good solutions have been suggested to discussing changing income distribution spread except the use of alternate measures, such as the correlation shown in equation 3.21 or rank-rank correlations which I will return to.

[14] Chetty et al. (2014) and Haider and Solon (2006) suggest that it happens earlier, by around age 30, in the US.

A final concern regarding IGE estimates is that they may not be linear across the parental income distribution. When estimating the standard IGE, one implicitly assumes that the logarithm of children's income is linear in the logarithm of parental income,

$$\mathbf{E}[Y_c|Y_p] = \mathbf{E}[\alpha + \beta_{IGE}Y_p + e] = \alpha + \beta_{IGE}Y_p \tag{3.22}$$

The third equality follows form $\mathbf{E}[e] = 0$ from inclusion of α in the model, and the fact that we condition on Y_p so that we can treat it as a constant when taking expectations. Under the linearity assumption, β_{IGE} does not vary with parental income. However, in the model above and in figure A.2 I showed that if just one parameter varies with parental income, this can create non-linearities in the IGE estimate. Becker and Tomes (1986) in addition suggests that credit constraints can create non-linearities. In applied work, many studies have found non-linear IGEs, including Bratsberg et al. (2007a) who show that Scandinavian earnings IGEs appear convex (while UK and US IGEs appear linear), Chetty et al. (2014) who shows non-linearities in US total income IGEs, Munk et al. (2016) showing a concave relationship in total income in Denmark using spline regressions, and finally Landersø and Heckman (2017), who also show a concave total income IGE for Denmark using local linear regressions. The fact that the IGEs are not linear is not an estimation problem in and of itself. However, it does mean that estimation of IGEs that do not take into account the non-linear nature may misrepresent actual mobility. As a result, if we wish to compare mobility across regions or countries, we must consider these non-linearities.

Rank-rank Correlations

An alternative measure of mobility is the the *rank-rank correlation*, which has been used recently for example by Chetty et al. (2014) and Boserup et al. (2014) to provide measures of intergenerational mobility in income in the US and wealth in Denmark, respectively.[15] The rank-rank correlation can be obtained as the coefficient from regressing the rank of children's

[15] The use of rank correlations to investigate aspects of inequality is not new.

position in their income distribution, R_c, on parents' rank in their income distribution, R_p.

$$R_c = \alpha_{RR} + \beta_{RR} R_p + \epsilon \tag{3.23}$$

The rank of both children and parents will lie in the interval [0, 1] and follow a uniform distribution. As a result, the rank-rank coefficient is equal to the rank-rank correlation (RRC), $\beta_{R_c R_p} = \rho_{R_c R_p}$, as child and parent distribution variances are identical at $var[R] = \frac{1}{12}(0+1)^2$. In the analysis I investigate whether expected child rank is linear in parental rank. A non-parametric investigation shows that this is indeed the case. While the coefficient provides an estimate of relative mobility, under linearity Chetty et al. (2014) argues that $R_c = \beta_{R_c,R_p,0}$ provides an approximation to an *absolute* measure of mobility. Given a parental income percentile, we can find the expected income rank of the child. More recently, Chetty (2016) and Chetty and Hendren (2016) have used these estimates to quantify expected incomes of children across commuting zones to show neighborhoods can impact the income prospects of children. I will return to these studies in later sections when I discuss how to quantify the impacts of neighborhoods on children in Denmark.

Relatively little is known about the potential problems of estimating RRCs in relation to intergenerational mobility studies. A recent study by Nybom and Stuhler (2016) investigates the properties of the rank-rank estimate (together with the IGE and transition matrix or copula). They find that also RRCs can be downward biased due to measurement error for parents' income.They suggest the solution also proposed for the IGE, namely averaging over several years when measuring parental incomes. They similarly show that the rank-rank measure reacts to life-cycle characteristics, although to a smaller extend than the IGE. Naturally, using too young individuals can result in highly misleading estimates comparing to later life if heterogeneity in income gradients over age is strongly heterogeneous. As a result, they suggest the same rule of thumb as for the IGE, namely to estimate RRCs around points of the age distributions that best approximates permanent earnings, namely mid-life. Problems of dispersions do not plague this estimator due to the uniform distributions of the two variables. Based on the estimates of Chetty et al. (2014) for the US and Boserup et al. (2014) for Danish wealth distributions, we might expect that children's expected income rank is linear in parental income rank. I will investigate this further in the following sections.

One questionable feature of the RRC estimator of mobility is that it removes the discrepancies in income at different percentiles - there may, for example, be a larger distance from moving from the 90th to the 95th percentile than in moving from the 30th to the 35th. This means that the measure does not quantify potential differences in welfare (if considered based on income as suggested in the model explained above). However, under the Rawlsian theory of justice, this difference may actually be morally relevant to consider (this is the difference principle).

Like the IGE, the RRC relates to intergenerational mobility through its inverse: The higher the RRC, the lower the expected equality of opportunity. The RRC is, however, less clearly linked to equality of opportunity than the IGE, for which we have developed an explicit model. Instead, we must make assumptions on the relations, suggesting that the RRC relates to the IGE through both depending on income. We can also consider directly the question of equality of opportunity through Roemer's algorithm: For each parental income percentage, do children asserting the same effort (that is, that are at the same relative level of performance in their type) gain similar outcomes? This naturally lead to the question of which characteristics we wish to equalize for, and to what extend they correlate with parental earnings.

The RRC is in essence a relative mobility measure - it tells us something about the statistical relationship between child and parental income ranks. It cannot tell us, on its own, where a child ends up in the income rank distributions (only whether expected outcomes of less advantaged kids is likely to be lower than those of advantaged kids). We can, however, construct a measure of absolute mobility using the estimated coefficients from the rank-rank regression in equation 3.23 (see e.g. Chetty et al. 2014, Landersø and Heckman 2017, and Rothbaum 2016). If we assume that children's expected income rank is linear in parental expected income rank, then the two coefficients α_{RR} and β_{RR}, together summarise the expected income rank of children conditional on parental income rank. This also means that if estimate the two coefficients, then we can predict the income rank of a child given a hypothetical income rank of the parent, such as e.g. the expected rank of children at the top income. This fact means that we can compare expected absolute upward mobility (for children at low parental income ranks) and absolute downward mobility (for children at high parental income ranks) between municipalities.

The justification of the absolute mobility measure hinges on the assumption of linearity. I therefore provide evidence of linearity of expected child and parental income ranks when showing baseline estimates of mobility at the national level. In the following subsection, I describe alternate measures of absolute mobility obtained from transition matrices. This is the measures that I will focus on in the empirical analysis in this section. I return to using the absolute mobility measure of expected child income rank based on the rank-rank regression in chapter 4.

Transition Matrices

A third recurring tool for measuring intergenerational mobility is the transition matrix. In essence, the transition matrix provides a joint probability density: Let a transition matrix be $A = ||a_{ij}||$, the cell a_ij gives the conditional probability that a child lands in group i, given that a parent is in group j (Richey and Rosburg, 2018). Let g_c be a mapping from a child's income to a child group and g_p similarly for parents, then

$$a_ij = \frac{Pr\left(g_c(Y_c) = i \text{ and } g_p(Y_p) = j\right)}{Pr(g_p(Y_p) = j))} \tag{3.24}$$

Where $\sum_i^m = 1$. An alternate description takes each cells value as the joint density, namely $a_ij = Pr\left(g_c(Y_c) = i \text{ and } g_p(Y_p) = j\right)$. This latter approach provides a link to copulas, a joint cumulative density over unit marginal interval (Nybom and Stuhler, 2016).

One important characteristic of the transition matrix compared to the IGE and RRC is that it provides information on upward and downward mobility - what share of children experiences upward mobility (and how far), and what share downward mobility conditional on parental income group (Black and Devereux, 2010). A drawback, as noticed by Mazumder (2011) and Corak et al. (2014) is that the group boundaries are arbitrarily drawn by the researcher.[16]

[16] They propose an alternative measure named a *Directional Rank Mobility Measure*, which gives probability that the rank of a child's income is above a threshold $s + \tau$, given that the parents income is at income rank s (Corak et al., 2014). The measure can be reversed to account for downward mobility.

In addition, unlike the IGE and RRC, the transition matrix does not provide a directly comparable measure of mobility (there is, so to speak no scalar value). Instead, an additional measure must be added on top.

The transition matrices are also affected by problems of life cycle bias like other measures. Nybom and Stuhler (2016) show that life cycle bias (measuring particularly parental income either too late or too early in the life cycle) can overstate mobility in the tails of the income distribution.

Several strands of literatures have provided different mobility measures M on transition matrices A (see e.g. Jäntti et al. 2006). Kanbur and Stiglitz (2016) provides a three part characterization for mobility (based on Shorrocks (1978)): We can define a perfectly immobile society by an identity matrix (children stay in the income group of their parents): $a_{ij} = 1$if $i = j$ and 0 otherwise, and a perfectly mobile society by complete row independence (parents group does not provide any meaningful information about child group). A third interesting transition matrix is a reversal matrix, for which the non-leading diagonal is 1. In this latter society, roles shift every period. Shorrocks (1978) offered a measure of mobility that lie in the interval [0, 1], give a value of 1 to the perfectly mobile society, and value of 0 to the perfectly immobile society. His measures is the trace measure (T)

$$T = \frac{n - Tr(M)}{n - 1} \tag{3.25}$$

Another alternative measure, suggested by Richey and Rosburg (2018), is a symmetrized second largest eigenvalue measure, which they assert is correspondent to correlations between generations within transition matrices.[17]:

$$M = 1 - |\lambda_2| \tag{3.26}$$

Where λ_2 is the second eigenvalue of the transition matrix A. This measure will become 1 if each cell takes the These measures of mobility allow for reducing the dimensionality of the transition matrices and making comparisons between different groups. To account

[17]The measure stems from an original article by Sommers, P. & Conlisk, J. (1979): "Eigenvalue immobility measures for Markow Chains". Journal of Mathematical Sociology, 6, 253-276.

for uncertainty in the estimation of the mobility indices, Jäntti et al. (2006) suggests using bootstrapped standard errors.

Hertz (2006) provides an example of the use of transition matrices to discuss differences in mobility between races in the US. He finds that blacks experience significantly stronger downward mobility when parents are middle income, and significantly less upward mobility from low income families than whites.[18]. This study is exemplary of the use of transition matrices to identify diverse movement patterns between subgroups in the total population. Likewise, Jäntti et al. (2006) investigates differences between countries using both transmission matrices and related mobility measures.

3.3 Other Studies in Intergenerational Mobility in Denmark

In order understand what we might expect in terms of the measurements of mobility, this section introduces the prior research on intergenerational mobility in Denmark. While none of the studies have emphasized spatial distributions of intergenerational mobility, they provide a baseline for the mobility estimates in the following section. The studies also provide a starting empirical starting point that I draw on in the next section where I introduce the data that I use for

While much research has been conducted on intergenerational income across Scandinavian nations, often comparing them to e.g. the UK or US, little research has been conducted in a Danish context compared to e.g. the UK or Sweden (see e.g. Blanden 2013; Black and Devereux 2010 for the more extensive literature not including Denmark, and Landersø and Heckman 2017 for a summary of Danish intergenerational mobility research). Danish intergenerational mobility has generally been found to be lower than in the US, and comparable (or potentially slightly lower) than other Scandinavian countries. Eriksson et al. (2005) provides first estimates of intergenerational mobility in Denmark using the Danish Longitudinal Study of Youth (DLSY). The panel followed 3,151 families and their children, born in 1954, over

[18]The findings have recently been corroborated by Chetty et al. (2018) who use IRS administrative data from the US to investigate absolute intergenerational mobility by race in the US.

time from 1968. As the follow-ups were provided not yearly, but in interviews with varying frequency (1970, 1971, 1973, 1976, 1992, and 2001), observations of earnings are limited. In addition, earnings are not measured continuously for fathers (reported in 1968) and children (2001), but in binned intervals; 7 for fathers, and 11 for the children. Given the above considerations of measurement error, the added binned data should make us more careful about drawing conclusions from this study. The estimate IGE's are at 0.29 for sons and 0.27 for daughters.[19] Sample attrition may also be a problem in the study (249 out of the 3,151 original children were not part of the 2001 interview), as may sample representativeness.

Jäntti et al. (2006) includes Denmark together with Finland, Norway, Sweden, UK, and US in an analyses with transition matrices. They find that Denmark generally is somewhere near other Scandinavian countries in terms of mobility, but they also conclude that the results using the mobility measures on the transmission matrices generally are uncertain. Clearer results arise from comparison of extreme movements. They find that Denmark both have the smallest probability of staying the poorest quintile given that father's are in the lowest quintile for sons. Denmark also have the highest probability of a son ending up in the highest income group given the father's income is in the lowest quintile. More interestingly, the same holds true in terms of son's of top quintile fathers entering the lowest quintile. The results are more mixed for daughters (with the US standing out in the comparison with less extreme mobility and more persistence).

Bratsberg et al. (2007b) include Denmark among a group of countries counting Norway, Finland, UK, and US, to investigate non-linearities in IGE estimates for fathers and sons. Their main motivation for the assumption of non-linearities is one of credit constraints, following the modelling by (Becker and Tomes, 1986)In the study they limit data from other countries to match that of the most restrictive dataset (UK) as best as possible. For the Scandinavian countries they find convex non-linear patterns, using higher order polynomial expressions, where lower than median income fathers experience an almost constant IGE, whereas for higher income levels, the IGE increases. This suggests that low income families experience

[19]Eriksson et al. (2005) particularly investigates questions of whether health may be a mediating factor in explaining intergenerational mobilities (they find that their IGE estimates are reduced by 28 and 25 percent for sons and daughters when controlling for various illnesses) it is still questionable to what extend reverse causality may be driving the results (although some studies are cited in their section 2 to argue the contrary).

more mobility, whereas families from higher income groups. In comparison, they suggest that using their sample of US data, namely PSID data, the IGE's in the US are closer to experiencing linearity. The authors show robustness of the Scandinavian findings towards standard issues in the literature, including measurement error and life-cycle bias is also proven. On a similar vein, Munk et al. (2016) investigates non-linearities in Danish intergenerational mobility. They also use register data, pairing classical IGE estimates with spline-regression. The underlying sample is fathers and children, with children aged 35-42 years in 2008. Father's income was measured in the years 1980-194, when the children were between 7-11 and 14-19 years old. They measure income as total income from earnings, own business, and capital income. Unlike Bratsberg et al. (2007b) who finds a convex relationship with earnings data, this broader income definition shows IGE estimates that appear concave, rather than convex. They also find a concave relationship from earnings, using oly the earnings element of income. Differences may be attributable to sample selections, and the inclusion of daughters in the final sample. At the top income percentile, they find capital income to be a main driver of persistence in positions.

Most recently Landersø and Heckman (2017) investigated income and educational mobility, comparing Denmark and with the US. The main finding of their income mobility analysis is that income mobility varies between Denmark and the US with income measures, the results indicating that Denmark is found to be more mobile due to a compressed income distribution.[20] Their analysis uses IGE's, as well as rank-rank correlations of children and parents' incomes in Denmark to point to the importance of income measures for determining

[20]In their analysis they include all children from cohorts born between 1973-1975, discarding those where (1) the child migrates, (2) the parent migrate, (3) there is no link to either a father or mother[21], or who have negative incomes across the full averaged period. The sample of parents-child observations is 166,359. Those with positive, averaged incomes in both generations make up 149,190 individuals. They include five different income measures (Landersø and Heckman, 2016b, Table A23): *Total gross income*, which includes wages, public transfers, profits from own business, capital income, and foreign income. *Total gross income excl. public transfers*, *net-of-tax total income incl. public transfers*, which is total gross income minus final income taxes paid, *wage earnings*, which include wage income and fringes, non-taxable income, severance pay, and stock options. Finally, the last income measure is *Wage earnings plus transfers*. The five measures of income is used to document that substantial differences in measured mobility arise from changing the underlying measure of income. For the US they use data from the PSID, having a total of 621 child-parent observations. Income measures are similar except no information about tax-payments exist in the PSID. When measuring parental income they use a 9-year average of income from over the child ages 7-15. For the Danish case, child incomes are measured in the years 2010-2012 (with the cohorts being 35-37, 36-38, and 37-39), for the US case where they have 621 observations based on similar restrictions, child incomes are measured in the last year of the five-year intervals of each of the included cohorts, 1972-1978.

the strength of relationships between parents and children. They find that when measuring gross income excl. transfers, the Danish IGE is at 0.352 (s.e. = 0.004), while the US IGE is at 0.312. They cannot reject a hypothesis that the two estimates are identical. When subtracting transfers from the total income measure, however, they find that the Danish IGE falls to 0.221. Including public transfers, they find that the Danish IGE falls to 0.271, while the US measure rises to 0.446. For wage earnings a stronger tendency appears: the Danish IGE falls to 0.083 (0.063 with transfers), while the US measure remains at 0.289 (0.419). As they note, comparisons across countries are hard to make (especially given the use of full registers in Denmark, and the PSID in the US), but the internal variation can be still be relevant to consider. They also take on a discussion of the importance of inequality, or the spread of income across generations in estimating parental incomes. They find that the extend of income inequality in Denmark is lower[22], and that Denmark has experienced less intergenerational increase in measured inequality.

The papers introduced in this section carry two main points. Firstly, intergenerational mobility in Denmark is generally lower than in other countries, and in particular in the US. Secondly, care must be taken in how mobility is estimated, and in particular in how the relevant income and family samples are created.

3.4 Data and Measures

Most recent studies of intergenerational mobility have turned to using administrative data sources as opposed to multi-generational sampled data. I utilize Danish administrative registers to investigate mobility at both the national and regional level. The register data has been made available through the IKE research group at Aalborg University, provided by Statistics Denmark. As access to the registers are limited on a per project basis, particular limitations have been put on the analysis. The data spans a period of 1980-2012 and include information such as age, family identifiers (including id of mother and father), educational attainment and enrollment, as well as income. In addition, the data contains information about municipal

[22] This is particularly attributed to a larger extend of wage compression in Denmark (Landersø and Heckman, 2016a).

location of most individuals in the Danish population in the same period. The data, however, does not allow for doing analysis at smaller geographical entities. As I will discuss in in the final chapters of the thesis, future research on topics of spatial variation in opportunity can gain from the additional information available in the complete register collections. In the following sections I discuss the data and measurement decisions.

The analysis in this chapter is centred on the national and municipal levels. For both analyses the main sample of the population is three cohorts of children born in 1973, 1974, and 1975, and their parents. After removing observations that cannot be linked to any location in Denmark, the total amount of children in the study is 205,625. Mobility measures are sensitive to both the age of measurement of income (*life-cycle bias*), as well as the length of the period (*measurement error*). Using the three cohorts from 1973-1975 and their parents provides a sample in which incomes parents can be measured over the longest possible period, and children's income at a period of their life-cycle where their relative income positions approximate permanent income.[23]

To measure the intergenerational mobility, I require estimates of parental income and children's income. Parental income is measured as the average CPI adjusted parental income while the child is between the age of 7 and 15 (the years 1980-1988, 1981-1989, and 1982-1990 for each cohort respectively). The income includes all income from salaries and personal business, public transfers, capital income, and non-classifiable income. The use of total income including public transfers provides a measure of total resources available in the family when the child was young. Alternate measures used in the literature includes labor earnings or earnings and transfers.[24] Measuring only earnings means restricting the analysis of interest to linkages between parents and children operating through labor markets, such as education investment and returns to education, parental social capital (Bingley et al. 2011; Putnam 2015, chapter 5), and transfer of non-cognitive social characteristics aiding in job requisition (Acemoglu and Autor 2009, chapter 1, Bowles and Gintis 2002). Using total income, including

[23] The choice of cohorts and measurement periods in addition follow that of other Danish research (e.g. Landersø and Heckman (2017)), which increases comparability of estimates from this thesis to other studies. A significant challenge in the international literature has been a lack of comparability of estimates (Solon, 2002).

[24] The most often used measure, due to its availability in panel surveys, is labor income (Black and Devereux, 2010).

government transfers and returns from financial holdings, incorporates all relevant income for determining socio-economic positions. Boserup et al. (2016) has shown that intergenerational wealth transmissions are substantial among high income families in Denmark, adding to differences in permanent income between the top 10 income families and the rest of the income distribution. For families living primarily off of wealth returns, neglecting this income, in addition, risks biasing the estimates of socio-economic mobility.[25] Related arguments apply for inclusion of social transfers, where exclusion of this source of income neglects the position improvement that transfer receiving families do not have.

Measuring parental income over a total of 8 years attempts to reduce measurement-error bias in mobility estimates (Solon, 1992). The choice of measuring income when the child is 7-15 is partly driven by available data. If we are interested in permanent income of parents, and the related averaged life-time socio-economic status, we would prefer to measure parental income over a period that better reflects life-time variation in income, rather than during adolescence of the child. However, this is partly not possible due to available time-frame of the dataset (1980-2012). A second, theoretical argument for measuring parental income during the ages of 7-15 is the interest of transmission of economic status. The Solon (2004) model assumes that parents invest in their children according to life-time permanent income. The model, however, makes strong assumptions of individual rationality and ability to predict life-time income.[26] In measuring parental income as the child is growing up, I get estimates of importance of family socio-economic status during childhood and adolescence for child outcomes.

[25]For a discussion of some of the theoretical arguments of relations between individual wealth distributions and individual economic outcomes, see see e.g. chapter 2 in Bowles (2012).

[26]This point relats to my discussion in footnote 10. For a parent to invest according to permanent income, the parent must form reasonably accurate estimates of permanent income (Y_p) based on total life-time income profile over a life-time of T periods ($Y_p^{Tot} = \sum_{age=0}^{T} Y_{p,age}$). The estimates of total life-time income must as such be accurate viewed from each age age t: $E[Y_p^{Tot}|Y_{p,0}, \ldots, Y_{p,t}]$. In addition, the individual must be able to make investments (I_p) in their children that is a function solely of permanent income (Y_p), and independent of yearly income fluctuations conditional on permanent income ($I_t \perp\!\!\!\perp Y_{p,t}|Y_p$). Insights from behavioral economics on *behavioral life cycle theory* emphasize limited self-control, mental accounting, and framing as potent determinants of household savings behavior, which will affect families opportunities to invest in children (Shefrin and Thaler, 1988). Recently Feldman (2010) has shown that families increase consumption as opposed to savings when receiving income in smaller amounts as opposed to larger lump sums, suggesting non-rational, but adaptive behavior in relation to mental accounts, and Thakral and Tô (2017) shows that taxi drivers act strongly adapts their economic actions in accordance with adaptive reference points updated at high frequence - in contrast with the expectance under the neo-classical permanent income hypothesis. The differences in assumptions and behavioral outcomes from the classic models (such as Becker and Tomes, 1979)) and those from the expanding field of behavioral economics (Thaler, 2016) suggests a new research in the intersection of behavioral economics intergenerational mobility. I do not expand on this topic here, but leave it as a suggestion for future research topics.

The measure of mobility may in addition depend on how parental resources are defined - either as father's, mother's, or total household income. The initial research on intergenerational mobility emphasized father-son income relations due to data availability (Solon, 2002) whereas recent literature (e.g. Chetty et al. 2014) has tended to emphasize household income as a more cogent measure of resources available in the household. In the results section I show that the choice of parental income (mother, father, household), affects estimates of mobility when measured at the national level. The findings suggests that care should be taken in comparing estimates of mobility emphasizing parental household income and individual parent income. As my primary interest is in estimating transmission of socio-economic status, I use parental household income measures in the remaining part of the analysis. If the child only lives with one of the indicated parents, the household income is only measured for that parent.[27]

Children's income is measured as total income in the years 2010-2012 for every cohort.[28] The income is measured when the children are 35-37, 36-38, and 37-39 years old for the 1975, 1974, and 1973 cohorts, respectively. As noted by (Solon, 2002), measuring children's income in the early stages of their career can severely underestimate the IGE (and potentially RRC, Nybom and Stuhler 2016), due to life-cycle bias from e.g. education pursuit. Landersø and Heckman (2017) show that IGE's and RRC's become stable for Danish income data in the second half of the 30's. This suggests that measuring children's income over this period provides stable and comparable IGE estimates for the three cohorts. Measuring the income in the same years but at different age levels will mean comparing individuals at different stages of their income life-cycle. However, measuring income in earlier years would require estimating incomes during

[27]It is possible that parents contribute to the social status of the child regardless of not living in the same household. However, measuring income in this way allows for inclusion of one-parent families. It is also possible that the income of other adults in the household may be important for determining child outcomes. As I cannot discern the share of income going to the child from parents not in the same household, nor the extend to which adults in the household that are not parents to the child provide additional relevant resources, I exclude them from analysis here. Another approach, taking e.g. by Landersø and Heckman (2017), is to measure parental income as the sum of parent and child incomes regardless of where the child lives, and discarding families in which there is only one parent. As this approach leaves out one-parent families, a potential at-risk group of families, I do not follow this approach. The construction of family income is one that merits further research.

[28] When measuring child income, I emphasize individual child incomes. To the extend that marrital status and partner income affects individual income (for example through choices of hours worked and taken child-care leave) focusing on individual incomes can, however, be misleading in tconsidering economic status. Eika et al. (2016) shows that marrital sorting (choosing partners within the same education level as one self) is substantial in both the US, UK, Norway, and Denmark. Properly assessing the importance marital sorting for intergenerational mobility in this context would, however, require additional research. I leave this for future research and here emphasize individual outcomes.

different stages of the Financial Crisis, starting in 2008. The stable income profiles in the late 30's, and the potentially strong impacts of the financial crisis through lost labor income at unemployment and financial return declines, gives favor to measuring incomes in the same years, at slightly different age profiles.

I construct income ranks for both parents and children using all observations *within cohorts*. That is, each parental rank is relative to other parents in the same child cohort, and similarly for children. Making the comparisons within cohorts minimizes biases due to different life-cycle income profiles and measurement periods.[29]

A persistent problem in investigating intergenerational mobility is how to handle non-positive income. Averaged parental household or child income is non-positive for 0.33 percent of the observations. Figure 3.1 shows densities of total and log income, for parents and children. Non-positive income has been transformed to take a value of 1000 DKK. [30] in particular show the presence of recoded negative and zero incomes on the left sides of each of the four figures. A non-positive income implies that the IGE cannot be estimated with the observations (as it requires taking the logarithm of incomes). More substantially, negative incomes arise due to personal business losses. It is reasonable to expect that business owners differ substantially from other parents either on public transfers or with limited labor market attachment. Former studies have handled this by either dismissing observations with non-positive incomes, or transforming the incomes to an arbitrary level near 0. I investigate the importance of excluding non-positive incomes in the following section for national mobility measures. I find that Exclusion as opposed to recoding negative incomes to 1000 DKK affects IGE estimates. This is less so the case for RRC measures. In accordance with this finding, I use all observations for income rank based regressions.

Table 3.1 contains summary measures for the CPI adjusted, averaged total income of individuals in the sample, as well as information on samples sizes.[31] Children, for whom I measure income at the individual and not household level, have average earnings in the sample of

[29] Income ranks are inherently relative, as opposed to income, which is an absolute income measure. The potential bias from measuring income at different ages, as a result, is larger for income ranks than incomes.

[30] Figure B.1 in appendix B.1 shows density plots without recoded zero and negative income, with blue lines indicating medians.

[31] Table B.1 in appendix B.3 contains a decile distribution of income for the samples groups.

about 396,865 DKK, just below that of fathers at 405,954 DKK. Mothers earn substantially less on average at 214,844 DKK. This is true across the income distribution as well. Parental household averages approximate the sum of averages of mothers' and fathers' incomes at 612,122 DKK, approximately 216,000 DKK above individual child income.

Table 3.1: Total income and sample sizes for children's, parental households', mothers', and fathers' total income

Statistic	N	Mean	St. Dev.	Pctl(25)	Pctl(75)
Child's total income	205,625	396,865	406,760	281,312	457,489
Parental household total income	205,625	612,122	350,064	468,285	693,152
Mother's total income	204,702	214,844	114,810	154,984	268,729
Father's total income	201,714	405,964	321,317	289,733	453,824

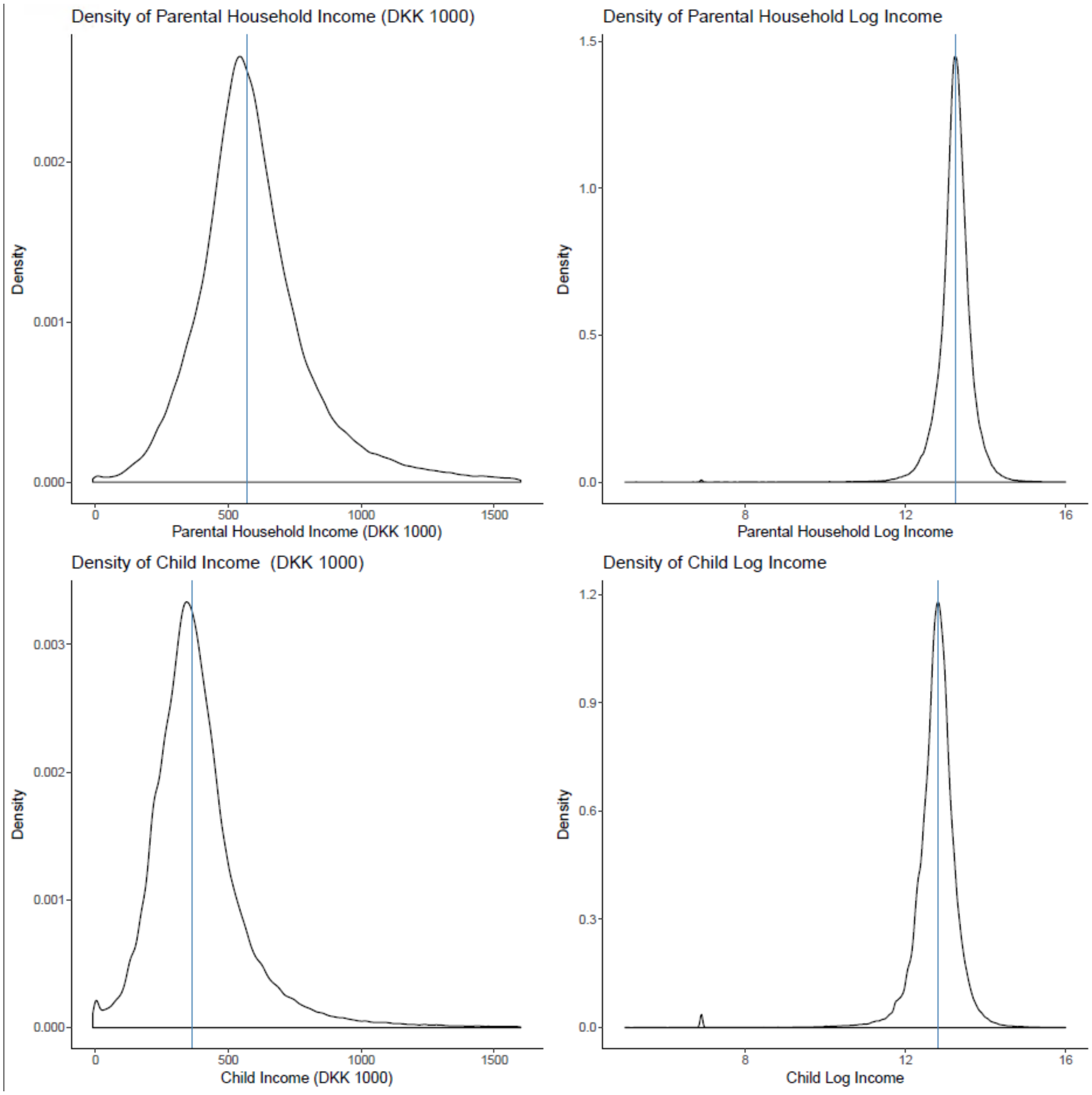
Note: The table contains information on the total income measure of all individuals in the main sample used in chapter 3. Total income is the sum of labor income, personal business income, income from financial wealth, and all transfers.

To investigate the spatial differences in intergenerational moblity, I use the spatial level of municipalities. Estimation of mobility requires a sufficient mass of observation in each entity, that is representative of the population in the municipality.[32] While studying mobility at lower levels of aggregation, these would likely not provide suffucient support with my available data. As a result, I am confining the analysis to the municipal level. In 2007 Denmark underwent a structural reform, where 278 municipalities were merged into 98 (and 13 counties into 5 regions). In the next chapter I turn to further municipal level analyses. Some of the data used in this analysis is only available for post-reform municipalities. I therefore merge pre-reform municipalities into post-reform entities. When estimating municipal level mobility, I assign each child to the municipality in which the child resided for the longest period between the ages of 7 to 15.

A natural concern regarding the division by municipalities is that this is an administrative, and not substantial division in terms of e.g. economic activity, or social belonging. This is a substantial concern (and one I that will return in the chapter on neighborhood effects), but given the available data this has been the smallest available entity of investigation (closest to neighborhoods). The results from the municipal level analysis can be viewed as a first

[32]The two main mobility measures, the IGE and RRC, are regression based. Lacking support in either the dependent or explanatory variables can lead to biased estimates if the lack of support is non-random (but even random variation can increase uncertainty in measures).

Figure 3.1: Density plots of parental household and child total income and log of total income.



Note: For each income density figure the blue line indicates the median. The upper left figure shows the density of total parental income in 2015 (in 1000 DKK). Total income includes labor earnings, income from independent business, capital income, and public transfers. Parental income is measured when the child is at the age of 7-15 years old. The bump at DKK 1000 is caused by recoded negative or zero incomes. The median household income is just above DKK 550,000. The lower left figure shows similar total income for children measured between the ages 35-37, 36-38, and 37-39 for children born in 1975, 1974, and 1973, respectively. A similar spike in the density can be found at DKK 1000, which again indicates recoded negative or zero income. The upper right figure shows log of total parental income, and and the lower left figure shows log of child total income.

approximation to be supplemented with further analysis.

3.5 Findings - The National Level

The IGE is the baseline mobility measure in most of the intergenerational mobility literature. It is natural to take it as the baseline for an analysis of mobility. The RRC estimate has become increasingly important, partly due to the obsevation that child income ranks often are linear in expectation in parental ranks in both the US and Denmark Chetty et al. (2014); Boserup et al. (2014). In this section I investigate the national level characteristics of mobility in Denmark and discuss potential problems with the IGE and the relevance of the rank-rank mobility estimate for analysis at the municipal level. I also emphasize the importance of measuring parental income at the household level.

I first estimate IGE's for the total sample with only observations with non-negative income, with observations with non-positive income transformed to positive 1000 DKK values. I estimate the results separately for household, father's, and mother's income. I also estimate rank-rank estimates using all observations, similarly for household, father's and mother's income. The results of the initial, national level regressions can be found in table 3.2.

Table 3.2: National Intergenerational Mobility Results

	Household	Father	Mother
Non-Negative Income IGE	.251 (.003)	.067 (.001)	.049 (.002)
Transformed Negative Income IGE	.230 (.003)	.069 (.002)	.049 (.002)
Rank-rank Estimate	.257 (.002)	.240 (.002)	.121 (.002)

Note: OLS standard errors are shown in parentheses. Non-negative income estimates includes all observations for which there is positive incomes. Transformed Negative Income estimates transforms any non-positive incomes to 1000 DKK. The rank-rank estimate is based on all available observations.

Table 3.2 shows that the average household level IGE estimate is 0.251 for non-positive observations, and 0.23 for all observations with transformed negative observations (these findings are close to those found by Landersø and Heckman 2017). For father's and mother's income separately, the IGE estimates are substantially lower 0.067 and 0.049 for father's and mother's respectively, using only positive observations. The result is almost similar for transformed negative incomes. The estimate goes up slightly for father's by 0.02 when recoding any missing observations. Mother's estimates are not affected. The rather substantial differences in mobility estimates suggests that looking only at only father's or mother's income may underestimate the coefficient (overestimate mobility), if we are interested in family resources. The small differences in estimates from recoding negative income in this case, suggests that we should be less worried about bias due to negative observations.[33] The size of these estimates are somewhat comparable to former estimates for Danish intergenerational mobility, where the smallest estimates lie around 0.05 (Landersø and Heckman, 2017; Hussain et al., 2009).

RRC estimates have a different pattern. The household estimate is at 0.257, whereas the estimates for father's is 0.24 and for mother's 0.121. As the ranks are made within the group (household, father, mother). The RRC estimates do not experience similar problems with missing observations. It appears that father's and household level mobility in ranks are comparable at national levels. In opposition, persistence in income ranks from mother to child is about half of that of Father's or the household level. The finding of differences between mothers and fathers income rank and related estimates once again makes for caution in how to measure income across parental groups. In particular, using mother's income ranks may underestimate the transmission of socio-economic status.

Underestimation of household level mobility when using father's and mother's income (for the IGE, and mother's for the RRC), suggests that one should be cautious about making claims about mobility when measuring only one parents income (unless this is the total household income). As a result, it makes little sense for me to compare my findings in the following part of the thesis to estimates based on single-parent income measures.

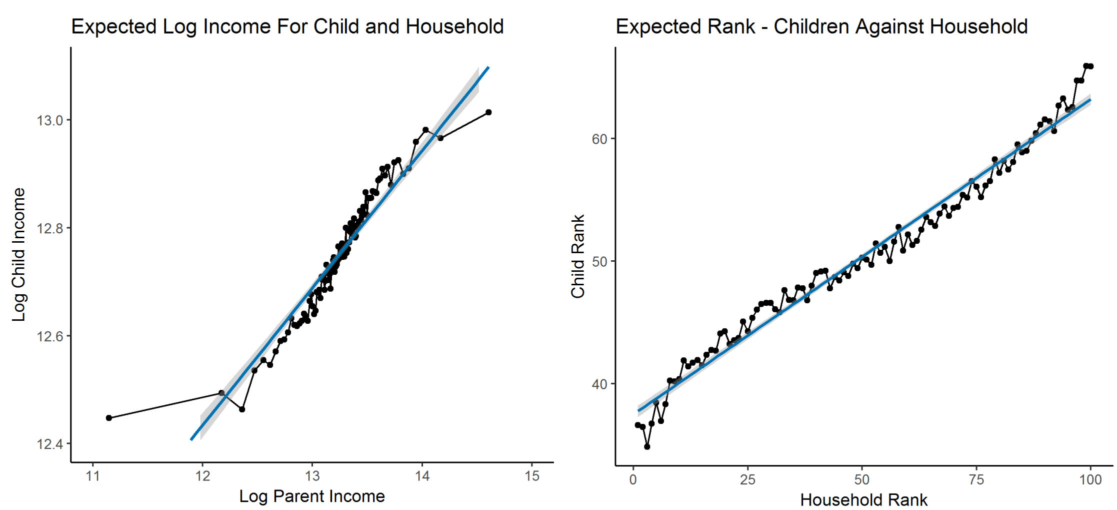
The former comparison does not tell us whether either the IGE or RRC estimate is a better summary of intergenerational income mobility. Both are linear slope estimates, but as discussed in the measurement section, non-linearities have been found for IGE estimates in a Scandinavian context. To investigate potential non-linearities in the log-income relationship I follow Chetty et al. (2014) and in the left side of figure 3.2 estimate the expected log child income given mean parental income for each of the 100 centiles of the log parental income distribution ($E[Y_c|Y_{p,centile}]$). This provides a non-parametric estimate of the relationship

[33] It suggests that the relatively small group of potential business owners with negative incomes are not skewing the estimates. The size of the group appears simply to be too small to impact the estimates.

between children and parents log income. If the IGE relationship is linear, then the expected value of log child income should be well described by the blue linear regression line seen in the figure. The relationship seems to be non-linear with a flatter (higher mobility) income relationship in the lower end of the distribution, and a larger slope (lower mobility) in the middle of the parental log income distribution.

The right hand side of figure 3.2 shows the mean child income rank given parental income rank for each centile in the parental income rank regression ($E[R_c|R_{p,centile}]$). The regression line fitted on the plot corresponds to the rank-rank correlation, $\rho_{R_cR_p} = \beta_{RR}$. It appears that the rank-rank estimates provice a slightly better linear fit than is the case for the IGE estimates, in particular for more extreme observations.[34]

Figure 3.2: Expected log income of children given mean income of parental household estimated by percentiles, and estimated rank of child total income given parental household rank.

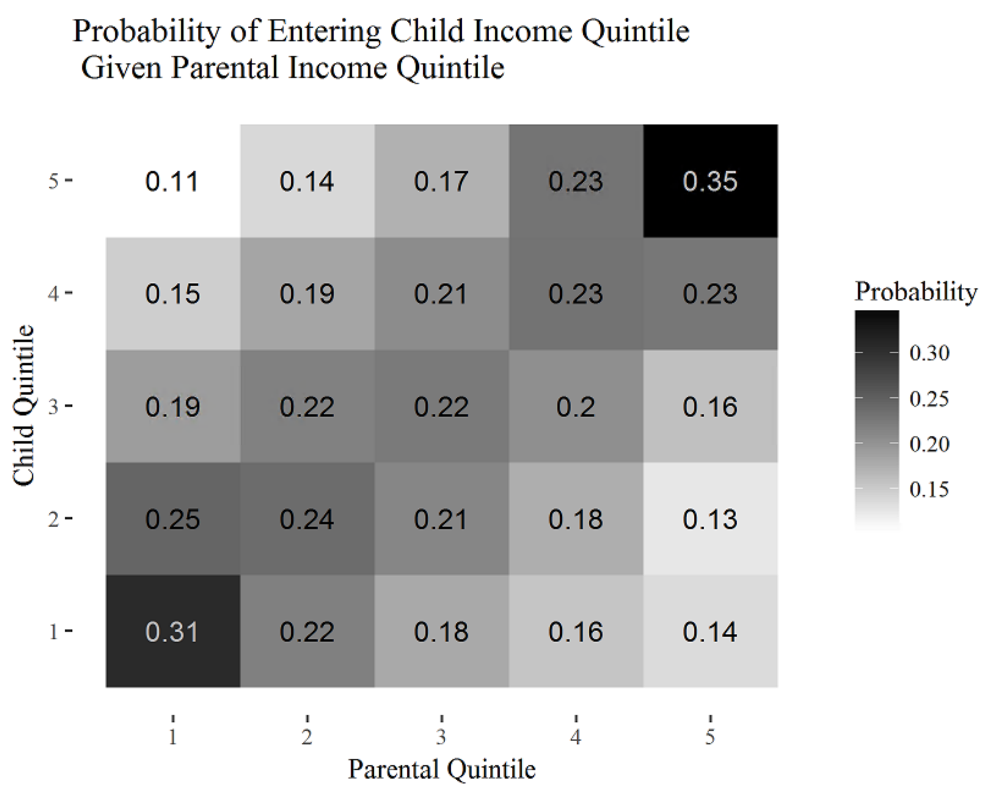


The rank-rank correlation appears to provide a better (linear) estimate of intergenerational mobility at the national level, and can be expected to also provide a better measure for investigating the municipal intergenerational mobility. In addition, as the rank measure is indifferent to non-positive incomes we can also include observations with zero and negative incomes in the samples when estimating the rank-rank coefficients.

[34]Tests for non-linear specifications, including inclusion of quadratic terms, and higher order terms could not be rejected at a 0.05 percent significance level. This is partly a result of the very large sample ($n = 205,625$). As a result, visual inspection of semi-parametric specification can aid in assessing the importance on non-linearities for the task at hand, which is to provide a simple measure of mobility.

The final measure of mobility in this section of national estimates is the transition matrix, and the probability of entering the top income quintile, when born in the bottom quintile. The matrix shows firstly that there is strong persistence in the ends of the income distributions. For children born in the bottom quintile of incomes, the probability of remaining there is 31 percent. For children born in the top income quintile, the corresponding quintile is 35 percent of remaining in the top quintile. This can partly be attributed to the impossibility of moving further down in income ranks from the bottom, and up from the top. Extreme mobility is less likely to occur - only 11 percent of the children born in the bottom income rank reaches the top income quntile, whereas for children in the top income ranks, the related probability of dropping from the top quntile to the bottom is 14 percent. Around the main diagonal movement up and down is more evenly distributed with the chance of entering quintiles just above or below largely similar to that of staying in the same income group between 20 and 24 percent. The overall picture is one of some mobility, with slightly less at the ends of the income distributions.

Figure 3.3: Transition Matrix with estimated probability of a child ending up in a quintile given the parental quintile.



In addition to the absolute mobilities apparent from the matrix, we can also calculate the Trace and M-measures of mobility for the matrix. These are 0.913 and 0.75 respectively for the matrix. As little intergenerational mobility research has used these measured of mobility, we cannot discern whether these are high or low values. Instead they will merely serve as benchmark values for the municipal level discussion.

With national baseline estimates of mobility, I next investigate differences in mobility measures across municipalities in the next section.

3.6 Mobility by Municipality

In this section I investigate the spatial variation in intergenerational mobility across the Danish municipalities. Like in the former section I initially estimate both the IGE and RRC. The estimates are based on OLS regressions for each municipality separately. Figure 3.4a shows a heat-map of the IGE's across Denmark, and the left-hand side of table 3.3 shows the municipalities with the 10 highest and lowest IGEs. The most immobile municipalities have measured mobilities similar to or above US levels (e.g. Chetty et al. 2014; Mazumder 2005, whereas the municipalities with the smallest IGEs have levels comparable to those found for father-son relations in earnings for Denmark (e.g. Munk et al. 2016 and Bratsberg et al. 2007a). From the map it becomes clear that mobility varies across the country.

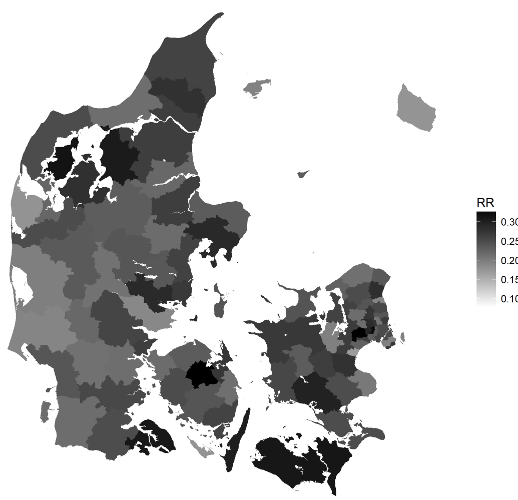
The rank-rank coefficients are shown in figure 3.4b, also in a heat map. A comparison of the two maps suggest that many areas have rather similar mobility estimates. The RRC estimates have smaller variance, and appears to be more evenly distributed across the country. Figure D.1 in the appendix shows densities plots of the two sets of mobility estimates. In the right hand side of table 3.3 I show the top and bottom municipalities in terms of RRC estimates. A comparison of the top and bottom mobility estimates suggests that the rank-rank relative mobility measure provides a less dispersed mobility measure. There is, however, a tendency for the same municipalities to be in the top and bottom of the distributions of mobility measures. The correlation between the IGE and RRC estimates is 0.68.

Table 3.3: Top 10 and Bottom 10 municipalities in IGE and Rank-Rank Coefficient Measured Mobility

IGE				Rank-Rank Coefficient			
Top 10		Bottom 10		Top 10		Bottom 10	
Municipality	Estimate	Municipality	Estimate	Municipality	Estimate	Municipality	Estimate
Fanø	.045	Albertslund	.371	Fanø	.099	Horsens	.292
Hørsholm	.144	Sorø	.373	Læsø	.168	Langeland	.294
Lemvig	.156	Kalundborg	.374	Lemvig	.175	Vesthimmerlands	.305
Bornholm	.160	Fredensborg	.378	Solrød	.180	Guldborgsund	.311
Lyngby-Taarbæk	.167	Sønderborg	.390	Bornholm	.186	Lolland	.312
Læsø	.174	Vordingborg	.400	Ærø	.191	Glostrup	.314
Solrød	.177	Hvidovre	.407	Lejre	.191	Sønderborg	.314
Billund	.186	Glostrup	.408	Hedensted	.199	Morsø	.321
Stevns	.200	Faxe	.420	Tårnby	.199	Høje Taastrup	.325
Hedensted	.205	Herlev	.427	Egedal	.202	Odense	.328

Figure 3.4: Heatmap of Danish Municipality IGE and Rank-Rank Estimates.

(a) IGE Estimates



(b) Rank-Rank Estimates

Note: The figures show Danish Municipalities and their (top) associated IGE for total child and parental household income and (bottom) Rank-rank regression coefficient estimate between child rank in cohort income group, and adult income rank in child cohort. The Island Bornholm has been moved from the far right to the upper right in the figure.

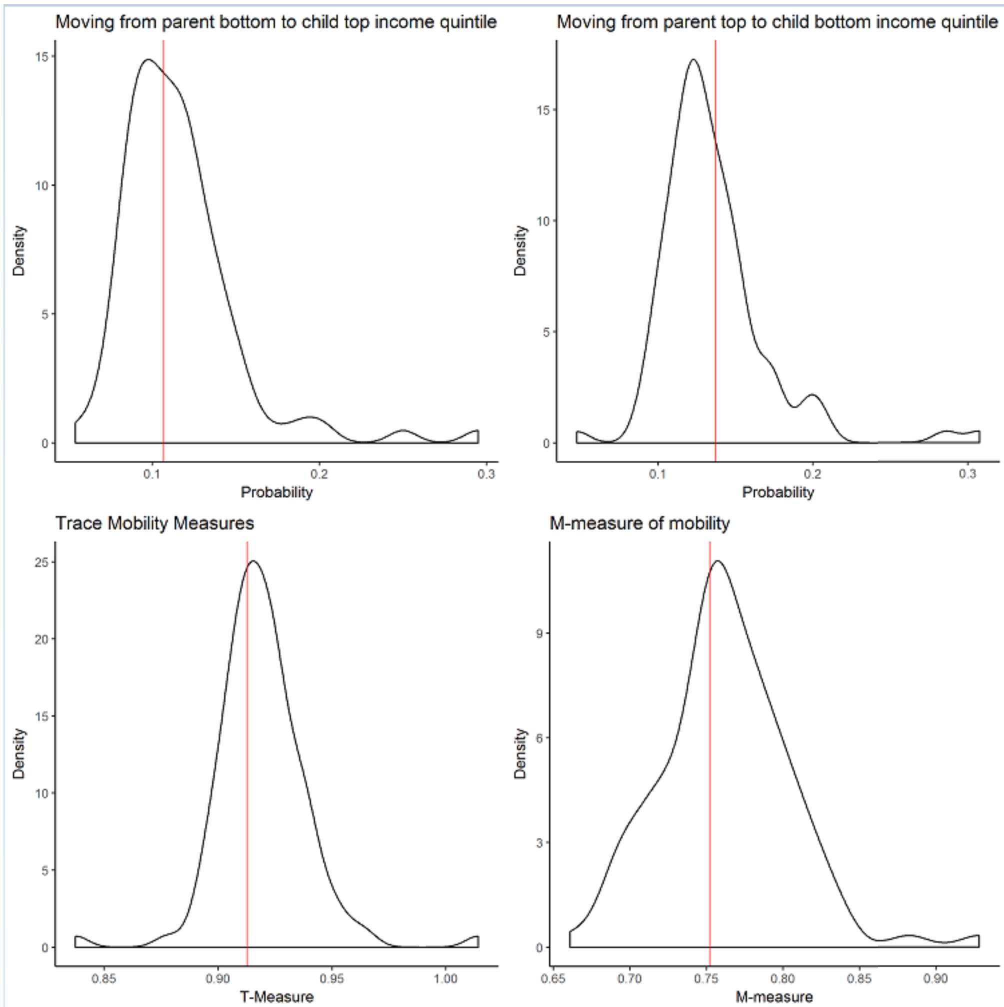
To get a sense of the scale of variation in the estimated coefficients, we can compare the estimated RRCs with estimates of county RRCs in the US by Chetty et al. (2014).[35] In appendix D.2, I show density plots of the RRC estimates obtained by Chetty et al. (2014). The US within country RR slopes differ from the Danish ones. This may partly be an attribute of different underlying samples of the population used for measurement[36]Chetty et al. (2014) estimate their RRCs using a sample of children born in 1980-1982. Their adult income is measured in 2010-2012, and parental income is measured in 1996-2000. This means that both the number of observations, and the underlying age-population differs from that used here. The small number of income measurements, as well as the early measurement of child income also should attenuate any estimates of mobility. The income definitions are comparable. Nevertheless, these estimates form a natural baseline in the literature given the large attention they have received in media and in academic circles (see e.g. Rothwell and Massey 2015b; Rothbaum 2016; Sharkey and Torrats-Espinosa 2017) More likely, they signify substantial differences in socio-economic mobility across spatial entities. The mean RRC for the US counties is .33, while for the Danish municipalities, it is 0.25. These differences in means correspond to what we would expect given national level RRC estimates (.25 for Denmark using the present sample, and .341 for the core sample in (Chetty et al., 2014, table 1)). The standard deviation of estimates is .04 for Denmark and .07 for the US. The estimated spatial variation in Danish RRCs is near half of that of the US.

To supplement the understanding of mobility, we can estimate mobility using transition matrices and their mobility estimates as in the former section. Figure 3.5 show the densities of the (1) the probability of a child moving from the bottom to the top income quintile - strong upward mobility, (2) the probability of a child moving from the top to the bottom income quintile - strong downward mobility, (3) the Trace-measure of mobility, and (4) M-measure of mobility. Red lines indicate national level estimates. Once again, it is clear that for the different mobility estimates, there are substantial variation in the mobility measures across municipalities. The standard deviations of each of the measures are comparable to that of the RRC at (1) .04, (2) .04, (3) .02, and (4) .04.

[35] The dataset has been downloaded from `http://www.equality-of-opportunity.org/data/` on April 25th, 2018.

[36] .

Figure 3.5: Transition Matrix Mobility Estimates for Danish Municipalities.



Note: From upper left (UL) to lower right (LR), the figures show densities of municipality estimates of mobilities based on transition matrices showing the probability of a child reaching an income quintile, given the parental income quintile, the child grew up in. (UL) The probability that a child born to parents in the bottom of the income distribution moves to the top of the income distribution. (UR) The probability that a child born to parents at the top of the income distribution moves to the bottom of the income distribution. (LL) The Trace-measure of mobility, T, calculated as $T = \frac{n-tr(matrix)}{n-1}$ for the conditional transition matrix, where n is the number of rows of the symmetric matrix, and tr is the trace of the matrix. (LL) The M-measure of mobility, calculated as $M = 1 - \lambda_2$, where λ_2 is the second eigenvalue of the conditional transtion matrix. Red lines indicate national level estimates.

For each of the measures that I consider, there appear to be substantial variation in the extend of intergenerational mobility. This, however, does not necessarily mean that each of the measures tap into similar estimates of mobility. In the measurement discussion I showed that each measure is dependent on the underlying measurement of income, that they could be affected by non-linearities, and for the IGE be dependent on spread of measured income in each generation[37], but also that they could be affected by non-linearities (IGE and RRC), as well as arbitrary groupings on income (for the conditional transition matrix measures). To get a sense of the extend to which the measures cover similar aspects of mobility, we can investigate the correlations of the different measures across municipalities. Table 3.4 shows the municipal level correlations of the measures.[38] The main finding in the correlation table is moderate relations between each of the measures. As noted above, the correlation between the most common measures of mobility is .61 at the municipal level. Both have medium negative correlations of -.43 and -.52 to strong upward mobility, but nearly no relation ($<$ .25) with strong downward mobility. The trace measure also appears only moderately related to the other measures, with the strongest relations with the RRC (-.5), and the M-measure (.63). The M-measure and the RRC are the two most correlated measures of mobility -.86.

These findings tell us that the factors driving the differences in the bottom to top movement across generations have little to do with the factors driving other differences. In addition, it appears that similar forces are driving the RRC, and IGE, and to some extend the M-measure of mobility. The strong relation between the RRC and M-measure is striking as the RRC is the preferred measure of the IGE and RRC (due to stronger potential for linearity). The M-measure in addition, is a measure of relations between cells in the conditional quintile transition matrix (Richey and Rosburg, 2018), using not only the information in the diagonal (persistence within quintile). The moderate relations to the strong upward mobility is, perhaps, less surprising. This is a limited measure of mobility, focusing on a particular group in the parental income distribution (the bottom fifth), and their somewhat unlikely strong upward mobility to the top.[39] A similar reasoning applies for the Trace measure of mobility,

[37]To the extend that not all income ranks are present in each municipalitity, this can also be an issue for the RRC measure in this case.

[38]The correlations at the municipal level cannot tell us about the underlying theoretical differences of the measures, nor what drives differences in them. It can merely give us an indication whether the factors driving differences in *the actual data* are similar.

[39]These cases are in and of themselves interesting in a Rawlsian perspective - how can we increase the

although this measure takes into account persistence within parentile income ranks, it cannot infer much about movements across different groups.

Table 3.4: Correlations of mobility estimates at the Municipal Level

	IGE	RRC	Bottom to Top Quintile	Top to Bottom Quintile	Trace Measure	M Measure
IGE	1					
RRC	.61	1				
Bottom to Top Quintile	-.43	-.52	1			
Top to Bottom Quintile	.01	.17	-.12	1		
Trace Measure	-.25	-.5	.26	-.17	1	
M-Measure	-.42	-.86	.42	.24	.64	1

In this section I have documented variation in intergenerational mobility estimates across municipalities. It appears that it does in fact matter where a child is born in determining the extend of equality of opportunity that the child experiences. The differences in mobility are present across a range of different measures. While there is some consistency in mobility across measures (see e.g. the maps in figure 3.4), the measures do not conform to provide a unified picture of mobility. While documenting differences in mobility in and of itself is a worthy goal (and the purpose of much of the research in intergenerational mobility), a more interesting question is why we find differences in mobility measures by municipality. In the next section I provide an overview of the main trends in investigating what drives differences in measured intergenerational mobility. An important emphasis is the difference between observational or data exploratory studies that investigate relations but cannot infer causality, and studies that attempt to estimate causal effects Tokey (1980). This discussion forms the transition to the next chapter, where I explore relations between municipal level covariates and in particular the RRC measure, as well as potential causal effects.

expected outcomes of the least advantaged in society? However, for discussing broader notions of mobility, including moderate upward mobility of children born in the bottom income deciles, it is a limited measure.

Chapter 4

Neighborhood Effects

In this section I introduce the theories on neighborhood effects, and studies that directly relate to the intergenerational mobility literature. The purpose of this overview is to construct a set of covariates that can be used in an analysis of the relations of municipal differences in mobility across municipalities that I documented in chapter 3. This part of the chapter is descriptive, but provides us with a first understanding of whether the observed mobility at all relate to variables that we would expect from the literature we are considering. Having shown that these covariates correlate not with relative mobility, but with absolute mobility (conditional on parental income rank), I then turn to investigating whether we can expect any effects of neighborhoods at the municipal level, developing first a causal inference framework, and estimating using families of movers between municipalities.

4.1 Theoretical Frameworks

The purpose of this section is to introduce the main theoretical constructs in the neighborhood effects literature in order to construct a set of potential predictors of the differences in mobility across municipalities that I documented in the former chapter.

Theories on neighborhoods and their effects have a strong history in the US, where issues of inner city poverty, especially among blacks, gained increased attention among sociologists after the publication of Wilson (1987), "The Truly Disadvantaged". In his book, Wilson characterizes black inner city poverty in Chicago, announcing the existence of what he calls the "underclass". The picture he paints is grim;[1] the inner city ghettos have high levels of unemployment due partly to decline of manufacturing in favour of service industry and relocation of production sites. Former inner city middle class blacks have left the old neighborhoods in favour of the suburbs, and as a result had withdrawn their resources and community engagement with them. The people in the inner cities lived in growing social isolation, losing access to job providing networks and mainstream societal values. The disadvantage in these societies were thought to have led to concentration effects that could partly be seen in mismatches between jobs and workers, low involvement in schools, too few marital partners, and little access to mainstream information sources such as networks and role models.

The main theories on neighborhood effects have been stated both by Wilson (1987) and by others after him (including Jencks and Mayer (1990)). The result is a broad collection of theories on how neighborhoods can impact individuals in them, and in particular the children in them. To characterizing the sociological theories on neighborhood effects, I draw on Small and Newman 2001; Leventhal and Brooks-Gunn 2000; Wodtke et al. 2016 from the sociological literature, and Oreopoulos (2003) from the economic literature to characterize the theoretical foundations in the literature on neighborhood effects.[2] The theories include *social isolation theories*, *social disorganization models*, *institutional resources*, *compound disadvantage theory*, and *relative deprivation theories*.

Social isolation theory derives from Wilson (1987). In essence, the theory states that growing up in deprived neighborhoods (especially the inner city neighborhoods), adults and children lack access to regular socialization and role models that can provide cultural messages, for example about the value of schooling, or how to act in certain circumstances. The socialization is (implicitly) expected to come from higher educated professionals, teachers, middle class

[1] An updated stance on his book is available in Wilson (1991). In the article, he restates the main theoretical framework that he draws on for the book.

[2] Small and Newman (2001) emphasized that there was a need to both develop the theories further and to take up empirical investigations to investigate the claims made by the theories. The fact that both Wodtke et al. (2016) and Chetty et al. (2014) refers to the same theoretical frameworks bearing on texts from the 1980's and 1990's suggests that theoretical developments were slow and that more careful investigations are needed.

individuals, etc. that are not present in the deprived inner city areas. The theory states that poorer areas with higher unemployment, few middle class individuals, and high unemployment will be characterized by little investment in education and human capital (or non-cognitive abilities) necessary to thrive. We can expect, as a result, that more deprived areas will have lower levels of mobility.

Social disorganization theory posits that in neighborhoods with lower degrees of collective social control, the extend to which members in the community take action to ensure order among (especially the younger) members of the community, there will likely be more crime. The higher levels of crime can affect the general emotional and behavioural development of children, leading to less education attainment. The effect are also expected to follow from less capacity in the neighborhood to supervise children, potentially leading to less education. The neighborhoods likely to be more affected are high poverty ones. The hypothesized effects from this theory is similar to those of the *social isolation theory*. There are no qualifications on mediation, time of effect, or other factors.

Institutional resource theories suggest that lacking resources, such as school quality, local libraries, after-school activities, etc. are likely to deprive children of opportunities. The theory suggest that large differences in available resources by area will give large discrepancies in outcomes. The emphasis is here on institutional resources, and not familial resources, or the immediate resources of other individuals living in the neighborhood. We can expect that government spending will be an important element in determining relative distributions of institutional resources in Denmark.

Compound disadvantage theory is a *glass half empty* framework. The baseline is a deprived neighborhood, in which higher income parents will choose to 'buy out' of the local area through private schools, extracurricular activities outside of the deprived area, or in other ways detach from the deprived social networks. The cultural context is the suspected negative effect (perhaps due to factors in the above theories). Children of the less well off are expected to develop 'anti' cultures that posits them against regular systematic experiences due to blocked opportunities and choices. Lack of role models and social networks from outside of the deprived neighborhood (as posited under the social isolation theory) can likewise induce less deprivation and subculture creation, including fatalistic attitudes. The lack of culturing can also impose less cultivation towards higher educational environments that children for higher income and education families might have. This effect will be aggravated, if the child spends relatively more time in the home, due to lacking opportunities in the local environment. Naturally, the compound disadvantage theory draws on the above theories in defining a within-environment outcome scenario. The theory implicitly posits that there will be a stronger transmission of parental income and educational status (through investment in the child) in more deprived areas.

Finally, *relative deprivation theory* takes a *glass half full* perspective. The theory posits that children from lower income families may not be able to gain from local neighborhood resources due to lacking family resources. Children may, for example, not be able to buy the equipment to participate in extracurricular activities, get high quality childcare, etc. This will be available for children from higher resource levels. A second mechanism in the theory comes from relative comparisons children may make to others (hence, relative deprivation) from which they may experience stigmatization from other children, or experience more negative self-perceptions that can impact possibility to participate in local activities. A final way in which this deprivation may occur is from competition for the available resources. If resources are scarce, and parental resources can influence children's abilities to make use of the resources, then children from relatively less well off families can end up falling behind. Implicitly the theory posits that intergenerational mobility will be stronger the higher the local costs of participation. In the opposite direction, in areas of lower cost of participation (generally less well off areas) there is a chance of more equal mobility levels due to less gains to reap for higher resources families.

The theories of this section have been developed with the purpose of explaining neighborhood effects in highly deprived inner city areas in the US. As a result, the theories may not be well suited for broader generalization on a national scale, nor even for understanding dynamics in cities such as Miami and Austin which differ substantially in racial composition, an important component in the original theoretical frameworks (Small and Newman, 2001). Naturally, the theories may be of even less predictive power in a Danish society, characterized by a more egalitarian income distribution.[3] In particular, if differences in income are attenuated by government transfers and taxation, then explanations of relative deprivation may not be relevant. If, in addition, government provision of services, such as child care, generally equal educational resources (including free tuition for higher education), this may further attenuate differences in mobility between neighborhoods under the compound deprivations theory. Naturally, local resources, and access to mainstream information through role models and networks, as well as local social control may still be important. Nevertheless, we may well expect that neighborhood effects will be less pronounced in Denmark.

The theories also gives a starting point in terms of what variables may be important in predicting differences in intergenerational mobility across neighborhoods. The concept of deprivation is persistent across the social isolation, and social disorganization theories. Wilson (1987) in particular focused on poverty rates, unemployment, and access to relevant role models. While potentially crude, measures of these characteristics are the share of the population living in *relative poverty*, the rate of *unemployment*, the share of *highly educated*, and share of high income earners (assuming that a larger share of these individuals raises the likelihood of them interacting with the children and having mentoring effects). Natural measures in relation to the institutional resources perspectives is the availability of libraries, the quality of local schools, and potentially the presence of education opportunities in the neighborhood. Unfortunately, I have not been able to collect information about the availability on such resources for this thesis. Ongoing work in the Center for Regional Dynamics at Aalborg University is presently mapping out the availability of such institutional resources. When this mapping is complete, it provides a highly relevant aspect to include in future research. The theory of relative deprivation suggests that with larger differences in parental resources, we can expect stronger relations between outcomes and parental background. This suggests the inclusion of *income inequality* in the neighborhood as another relevant measure. From both the social deprivation and disadvantage theories, cultural norms are important drivers of outcomes, for example through the extent of social control in the local area. One of the frequently occurring

[3] Landersø and Heckman (2017) provides distributional kernel density estimates of income distributions for Denmark and the US that investigate data similar to what I use in this study. Based on Danish register data and US Census data, they show that Denmark not only have more compressed income distributions (after transfers, before taxes), but also that Denmark does not have a substantial group of low income earners that exist in the US.

predictors of disadvantage is the share of *teenage mothers* and *teenage births* in the local area. *Local crime* is another variable that is of interest, especially following social disorganization theory. A final element that has received considerable attention is the extent of segregation of racial and ethnic types. With spatial aggregation limited to the municipal level, I cannot discern within-municipality segregation between e.g. ethnic categories. Recent discussions of ghettos in Danish media and politics suggests that this may, however, also be important in Denmark. As a first proxy, I therefore consider the share of *non-western* individuals.

Several variables and aspects have not been included in the above. Putnam (2015), for example documents that an important part of local area advantage and disadvantage relates to social capital. While measures of social capital exists for the US (see e.g. Chetty et al. 2014), I am unaware of similar spatial mappings of social capital in Denmark. This leaves me without the opportunity to investigate it. I also do not investigate within school environments, which Landersø and Heckman (2017) suggest may be an important aspect of discerning differing mobility in Denmark. Lacking the available register data to discern the characteristiscs of these local schools (e.g. local GPA's and teacher quality), I must also confer this work to future research. A third concern is the level of measurement of the above variables, which I will return to in data section 4.3. The theory is related to neighborhoods, with definitions of neighborhoods being variable. Unfortunately, investigating intergenerational mobility and neighborhood effects, I am limited to certain levels of spatial aggregation above for example inner city areas.

In the next section, I introduce the empirical findings on intergenerational mobility and neighborhood effects. A substantial concern when working with neighborhood effects is that any relation between neighborhood variables and intergenerational mobility might in fact be driven purely by parents self-selecting into specific areas. Surveying the existing evidence, I find substantial support for pursuing an investigation of the predictive capacity of the neighborhood variables that I have introduced in this section.

4.2 Studies in Neighborhood Effects and Intergenerational Mobility

Studies in the intersection of the literatures on neighborhood effects and intergenerational mobility are scarce relative to the many studies in the separate literatures. In this section I consider both the main concerns of estimation of neighborhood effects, and draw on empirical studies to exemplify the problems and solutions as well as the main findings on the effects of neighborhoods on intergenerational mobility.

A major concern in neighborhood effect studies is selection bias. When we move into a new neighborhood, we choose the neighborhood. This means that if we were interested in how being in a neighborhood affects the individual, we cannot simply compare means of outcomes from individuals in different neighborhoods. We would be risking that the individual living in Aalborg was different in a way from the individual in Gentofte such that we could not identify which was the effect of the individual and the neighborhood.[4] Naturally, children cannot themselves opt into neighborhoods (at least not when young). Instead we might be worried that there are e.g. parental characteristics effects that are predefined, and potentially could be attributed to the neighborhood effect. In essence, this would (in simple cases) be a problem of omitted variable bias.

Much of the intergenerational mobility literature has been emphasizing the importance of parental characteristics in determining child outcomes, either neglecting or finding evidence of minimal importance of neighborhoods, but some emphasis has been granted to attempts at bounding variation solely arising from factors external to the child and its family. The common framework for investigating the variation due to e.g. neighborhoods is decomposition of variance. Black and Devereux (2010) describes the main idea used on siblings as follows: Let ρ be the correlation between siblings in earnings (or another outcome). Then, under assumption of additivity, this correlation can be decomposed into $\rho = \beta_{IGC}^2 + s$, where β_{IGC} is the intergenerational correlation coefficient. This term approximates elements that

[4] The basic problem can be posed both in terms of unknown (and therefore) omitted variables in a regression framework, or as a problem of is a basic theme in identification in either a regression setting with omitted variables, or in the casual effects literature, see e.g. (Angrist and Pischke, 2009, p. 12).

are similar to siblings and correlates with parental characteristics. s are any other factors that both siblings share, but which are not correlated with the parental characteristic (Björklund, 2016). The extent to which factors such as neighborhood effects, parental income and networks effects, and more, are separable in this sense, and to what extent the findings are useful may be questionable. Nevertheless, this type of studies have been pursued until recently (see e.g. Björklund and Jäntti 2012; Björklund 2016). The main findings are that siblings correlations are stronger in the US (Mazumder 2008 finds correlations of 0.49 among brothers) and Germany (Schnitzlein 2014: 0.432 for brothers, 0.391 for sisters), and lower in the Scandinavian countries (Björklund et al. 2009: 0.32 for brothers in Sweden, Schnitzlein 2014: 0.202 for brothers and 0.187 for sisters in Denmark). With squared IGC's broadly similar to these correlations across countries, the general finding is that up to four fifths of the siblings correlations in earnings may be due to family characteristics (Black and Devereux, 2010). The remaining variation to which we might ascribe neighborhood effects appears small at best. In a related literature, Page and Solon (2003) investigates correlations between children from the same neighborhood, and siblings within the same neighborhood. They infer that for the US and using the Panel Study of Income Dynamics, the siblings correlation was higher than 0.5, while the correlation among neighbours instead was about 0.2, and potentially closer to 0.1. Raaum et al. (2006) do the same for Norway with administrative data. They find much lower estimates here (.05), leaving almost no variation purely to the neighborhood.

The large explanatory power of family background (often measured through income) relative to other factors begs the question of what it is that drives these differences - is it characteristics that are innate to the families? In this case, under both the Rawlsian and Roemerian approach to equality of opportunity, we might not want to consider this a violation of equality of opportunity. In addition, it suggests that little room is left for discussions of the impact of neighborhoods. A tool often used to consider the extends of importance of e.g. innate traits, but also such things as quality of schooling is decompositions of mobility estimates. Most decompositions, however, are regressions that tease out shares of variance that (under the model and functional form) approximates shares of explained variance. An important example with respect to the importance of innate characteristics is Bowles and Gintis (2002), who decompose parent child earnings correlations into additive direct and indirect effects of various factors.[5] They find little effect of cognitive heritable traits, whereas they emphasize (in line with Bowles and Gintis 2011, originally published in 1976) that other non-cognitive skills, such as the ability to adhere to routines and reward systems, are important determinants of labor market outcomes, and of intergenerational relations. In an updated version of the study (Bowles et al., 2005) emphasize the importance of schooling (conditional on IQ), and phenotypes, such as race for determining outcomes. Osborne Groves (2005), in a similar vein, investigates the importance of non-cognitive traits, such as fatalism, finding that family traits appear to be important in determining the extend of transmission of outcomes (for US data). Blanden et al. (2007) provides another decomposition for UK data. They find that non-cognitive skills, ability labor market experience, and employment interruptions can explain up to half of .32 IGE. Also for the UK, Blanden et al. (2014) suggests from a decomposition by mediation analysis that education level is a stronger mediator of educational outcomes than occupation level when compared to to the UK. As a final and most recent example, Björklund et al. (2017) investigate the importance of birth weight, birth height, grades, and final education for IGE estimates with data from Sweden and the UK. They find suggestive evidence that grades and final education are particularly strong mediators of parental outcomes, whereas birth weight and -height are not.

The findings of little importance of innate characteristics suggests that there is indeed room for neighborhood effects discussions. This is particularly the case when considering that parental behaviour may also be affected by the neighborhood they are in over the course of childhood and adolescence of the child. The early childhood development literature (summarized by Heckman 2008b) has provided ample evidence that phenotypical traits can be affected by environments, meaning that even genetic effects may be malleable. This means that even while the above shares of siblings correlations relating to parental income are large, this may be due to interactions between other environmental factors and parental income (or its correlates). A stronger interactive effect may also appear when considering parental behaviour that is also contingent on environment. Even while there is indications of some causal effects

[5] In order to obtain the decomposition, Bowles and Gintis (2002) make use of normalized variables, so that the coefficient of each variable in a regression onto child earnings can be understood as “a one standard deviation increase effect”. The statement that this is an actual effect may be unwarranted. There is no clear reason why the regressions they perform should provide causal estimates as opposed to descriptive ones that are contingent on the specified regression.

of e.g. schooling [6] and local industry changes and worker displacements (Oreopoulos et al., 2008), these effects may in fact be filtered through neighborhoods.

As noted by Small and Newman (2001) an extensive literature has made use of observational data to attempt to find the relative importance of neighborhoods for individual and family outcomes. Nevertheless, the observational data most often is susceptible to the problem of selection bias and observational problems. Harding (2003) suggests to use a propensity score matching framework to deal with estimation of the effect of neighborhoods. The propensity score matching estimator in essence attempts to find individuals in the dataset that are as alike as possible in (estimated) probability of treatment (Angrist and Pischke, 2009, pp. 86-91), thereby overcoming problems of specifying functional forms of regression model relationships. King and Nielsen (2016), however have recently suggested that propensity score matching may not be particularly helpful as it is likely that while estimated likelihood of treatment (neighborhood effect) may be the same, the two observations can be substantially different, making it plausible that (1) substantial noise is added to the estimation, or more severely (2) that the estimation is not plausible. The golden standard alternative to the methods implemented by Harding (2003), IV regressions, or other estimation techniques are *experiments*.

Across 5 major cities in the US[7] the U.S. Department of Housing and Urban Development (HUD) implemented a program to provide families from high poverty neighborhoods with an opportunity to find a place to live in lower poverty neighborhoods (poverty rates less than 10 percent). Gennetian et al. (2011) provides an overview of the experiment. The key feature of MTO that renders it highly useful for investigating neighborhood effects is the *randomization* of treatment. Families that participated in the program were assigned to one of

[6] Pekkarinen et al. (2009) use a differences-in-differences methodology to find the effects of decreased school tracking on intergenerational mobility in Finland, drawing in a reform that lead to later tracking of students.They find that the reform resulted in a drop of the IGE from .3 to .23. In another study, Mayer and Lopoo (2008) use state fixed effects in regression models to determine the importance of state-level spending for intergenerational mobility. Their identification of effects on changes in spending over time within-state. They find that changing government spending from $9,157 dollars per child to $16,807 (going from mean in lowest third of spending to that in highest third spending), an increase of 84 percent, leads to a drop of 34.6 percent in the measured income persistence - approximately 0.17 log points. This study, however, warrants some caution. Their suggested changes in spending is unlikely to appear within states (there is not support for the regression in this span), and it is likely that spending covaries with other changes, meaning that the effect is not identified.

[7] The cities were Baltimore, Boston, Chicago, Los Angeles, and New York.

three groups: (1) the experimental group, where participants got a voucher that would reduce their rent in a low poverty neighborhood, and guidance from counsellors both for finding a new place to live and in the new neighborhood. (2) A group that received only the voucher with no requirements.[8] (3) The control group that did not get a voucher. Participants were unaware of their assignments to groups. To be considered for the program, the participating families had to sign themselves up for participation. The MTO sample of families generally is not representative of the general families in the US.[9] 98 percent of the families were female headed, 66 percent African American, 29 percent Hispanic, and nearly none were white. Only 40 percent of parents had a high school diploma, and between 25 and 29 percent were working (74 to 75 percent were on Assistance to Families with Dependent Children (AFDC) welfare benefits).

The randomization ensures that the groups are mutually comparable, and no family can influence their own likelihood of receiving the treatment. As a result, the MTO provides selection-free estimates of neighborhood effects, and researchers have used it extensively to investigate effects on health, crime, welfare participation ((Leventhal and Brooks-Gunn, 2003; Ludwig et al., 2008; Gennetian et al., 2011), and most recently for long run child outcomes (Chetty et al., 2016b). The findings from the study have been mixed. For parents, no effects were found on jobs, whereas slight improvements of mental health and feelings of distress were found (this is consistent with the fact that a main reason for 77 percent of participants for moving was to get away from drugs and crimes, which actually happened). For children the outcomes were mixed. For males, math test scores generally became slightly worse for the experimental group than the control group and no difference in reading scores, while both became better for females. The males also committed more property crime, and were more likely to have used drugs. In the longer run, Chetty et al. (2016b) find that the effect of moving on children generally varies with the age of the child using administrative income data. For children below 13 before moving, the experimental group children had a mean of 31 percent higher wages in the mid 20's (the control groups mean was $11,270), and were more

[8] The voucher is called "Section 8" voucher, and is a generally utilized tool in US housing policy since 1978 (see NYCHA 2018). The voucher provides the receiving family with the opportunity to have a share of their rent covered (depending on income, and the adjusted price of rental).

[9] Sampson (2008) found that the distribution of characteristics of the families participating in MTO corresponded to about 5 percent of the families living in Chicago using a representative sample of families.

likely to have gone to college. For children 13 or older at the time when the family moved they find no effect of the better environment. It is noteworthy that these findings generally are in accordance with the early childhood effects of investment suggested by Cunha and Heckman (2007); Heckman (2008b).

The findings from the MTO experiment suggests that there may be long run gains from better neighborhoods. However, as noted by Sampson (2008), the MTO experiment probably is not generalizable. As noted, the sample of families are representative of about 5 percent of the Chicago population, a large US city. This suggest that we cannot necessarily extrapolate the results to general population, nor answer questions of interactions with intergenerational effects (note that the findings could seem to support the cumulative disadvantage theory). As a result, while the MTO sets a gold standard for how neighborhood effects potentially can be measured, other evidence is necessary as well to understand how, if at all, neighborhoods is one of the circumstances that we may wish as a society to remove the effect of in the name of equality of opportunity.

Oreopoulos (2003) provides an early study of the effect of neighborhoods on child earnings. He utilizes random assignment in projects (public housing) to different neighborhoods in Toronto, matched with administrative income information. He finds no effect on earnings and education of children who were assigned to different neighborhoods of public housing. While the study's methodology provides a related identification to that of the MTO, a similar problem of generalizability also arises. The children growing up in the projects generally cannot tell us about the broader population as these live in relative poverty. Sharkey (2008) provides an interesting addition, investigating intergenerational *context* mobility - the similarity of parents' and children's neighborhoods (measured using mean neighborhood income). He utilizes PSID data to estimate the model, finding a strong IGE estimate of intergenerational neighborhood attainments of 0.64. This is suggestive of strong persistence in neighborhoods over generations. Part of this may be attributable neighborhood effects. In comparison, individual IGE estimates for the US generally (with the exception of Mazumder (2005)) are smaller at about 0.4 (Black and Devereux, 2010). The study by Sacerdote (2007) also make attempt at estimating the effects of neighborhoods, using the random assignment to adoptive families of adoptees from South Korea. In a multiple regression framework he finds no effect on income of adoptive children from multiple neighborhood characteristics (after controlling for adoptive parents‘ characteristics), including zip code average income, and education, urban density, nor the share of black parents. This is a strong suggestive evidence for little actual neighborhood effects for a broader outlook of the population, although we may worry that especially Korean adoptees suffer for example discrimination in their local environments that makes the evidence non-generalizable.

Chetty et al. (2014) provides the first comprehensive study of the variation of equality of opportunity, or intergenerational mobility in the US, and while they do not estimate effects of neighborhoods in this study (they do in Chetty (2016); Chetty and Hendren (2016)), this study is essential by forming the background both for this thesis and for other research on neighborhood effects on intergenerational mobility. They are the first to utilize full population registers of the US population to estimate the textent of mobility in the US. However, limitations in the available data limits them to investigate samples of children born between 1980 and 1991, and measure child income as the average family income for the child and spouse over 2011-2012. They measure parental income at child age 15-20, and the child's income is measured up to age 32 using their core sample. Income measures are total pre-tax income, including transfers. To measure the relative mobility, they estimate IGE and rank-rank coefficients. More interestingly, however, they emphasize absolute mobility measured in several ways. Firstly, using transition matrices, they map out the chances of entering the top quintile with parents in the bottom quintile. Secondly, they estimate absolute mobility using their rank-rank estimation. As they find that child expected income rank is nearly linear in parental income rank, the estimate absolute mobility as the expected rank of a child given parental income rank (most often 25th percentile). Among their findings are large discrepancies between areas when evaluating the expected child outcome at the 25 parental income percentile (a difference of 10 percentage point between the top and bottom of the 50 largest commuting zones). They in addition show how the measured correlations vary with income inequality (negatively, -0.578), top 1 percent income share (-0.19), racial segregation (-0.361), school quality (with various measures, average of 0.6), social capital (0.562), and finally various measures of family structure (-0.76 for share of single parent families). Family structure is also the strongest predictor of future outcomes in their study. This is the study closest to this thesis. However, while Chetty et al. (2014) find large variability in their estimates for the US, there is reason to believe that the same will not hold true for Denmark.

Several studies have used the mobility estimates of Chetty et al. (2014) as they have become available online.[10] Among them are Sharkey and Torrats-Espinosa (2017) who uses an IV estimation approach to investigate how changes in crime levels may relate to mobility levels. They find indications that a one standard deviation crime reduction leads to 1.2 percentage point higher income for a child of a parent at the 25th percentile. Other studies are less rigorous, and hence plausible. Rothwell and Massey (2015a) provides OLS regressions to show that the estimates correlates with higher school quality.[11] Donnelly et al. (2017) follow a similar strategy, but acknowledge that they are investigating correlations and not causal effects. They regress children's cognitive test outcomes as well as externalizing behaviour test scores (from the Fragile Families and Child Well-being Study) on county mobility estimates (given parental income level), and control for a variety of measures. They find a 1 standard deviation increase in mobility is associated with 0.14 standard deviation increase in vocabulary scores, and .11 standard deviation in reading scores. Finally, Chetty (2016) and Chetty and Hendren (2016) uses the estimates from Chetty et al. (2014) to study the effects of neighborhood effects using moving families. They find that for each additional year spent in a 1 percentage point better neighborhood (measured as the expected income rank of a child from the corresponding environment), the expected outcome of a child increases by 0.04 percentage point. The full population coverage of the IRS administrative data makes this study the most compelling for investigating the effects of environments across the ones mentioned in the literature review. However, as I have suggested above, there is still reason to believe that the variability and effects follows from the institutional context.

The discussions in this section have several rammifications. Firstly, there is reason to believe that neighborhoods and their related characteristics do in fact impact child mobility. The variance decompositions in the neighborhood effects literature suggests that little variance is left

[10] `http://www.equality-of-opportunity.org/`

[11] This study is somewhat misguiding. The author's insist on calling their estimated coefficients for effects, while these cannot be attributed any causal interpretation. In addition, it appears that in their regressions, they may be controlling for variables that they should not.

to neighborhoods to explain in outcomes. These studies, however, make strong assumptions of additivitity of effects, which might not be warranted, in particular given the contingency of phenotypes on surroundings - an important element of the early childhood development literature. The studies on neighborhood effects and intergenerational mobility generally provide varied results. Most have been performed for subpopulations that are not representative of populations, and non have been performed in a non-US setting. The most recent strands of the literature provides estimates of neighborhood effects and differences in mobility estimates at the national level, all of which suggests some scope for the effects of neighborhoods, and it may be an interactive one. A general problem that arises whenever we study neighborhood effects, however, is the extent to which any relations between neighborhoods and mobility is driven purely by self-selection. While neighborhood effects may exist, this can make any observed differences in mobility across areas an artifact of family background. Nevertheless, in order to argue that some effect of neighborhoods exist, we should be able to discern statistical relationships between neighborhood covariates and intergenerational mobility across areas. This is the goal of section 4.4. In the next section, I describe the data used for the operationalization of the theories from section 4.1.

4.3 Data and Neighborhood versus Municipalities

The first step I take to investigate whether the observed differences in mobility can be explained by municipalities, and not by family characteristics is to look at the statistical relations. In the former section I noted several factors from the neighborhood effects theories and literature that may influence intergenerational mobility and implicitly equality of opportunity. In this section, I describe the covariates I include to look at correlations between the observed mobility, and neighborhood characteristics.

The neighborhood effects study optimally requires data at neighborhood level. In Wilson (1987) and many of the MTO studies, the neighborhood level is defined either using administrative boundaries or areas defined by characteristics. They are, however, mostly measured in smaller geographic entities - e.g. blocks, or city areas, than the data that I presently have available. The mobility estimates from the former section are measured at the municipal level due to the number of observations required for estimation, and the data that has been generously made available for the thesis. As a result, there is a discrepancy between what can be measured and the spatial level of the theories. Within municipalities there can be large differences between different areas,[12] but I must defer to later studies to investigate the importance of these at lower levels spatial disaggregation.

When constructing neighborhood covariates, I have been similarly constrained to the municipal level. I consider 7 main covariates in the correlation analysis: (1) the share of adults with above secondary education, (2) The share of working age (15-64) population that are unemployed, (3) the share of individuals with income in the national top decile, (4) the share of individuals with income in the bottom decile, (5) the share of non-western first and secondary immigrants, (6) The share of teenage mothers to population, (7) the number of teenage births to population, (8) the within municipality income inequality, and (9) the number of crimes reported to population. I measure each variable as the within municipality average over the years 1980-1990 when possible. The period overlaps with the period in which I measure parental income.

(1) High Ed: The share of adults with above secondary education is calculated using final education information of the total population above 18. The data is only available for 1981-1990. For a substantial share, around 30 percent across years, there is no education information. A large proportion of individuals with no registered highest education is the elderly. The share drops to approximately 18 percent when looking at the population below 64. High education individuals are defined as all individuals with registered professional bachelors or university degrees (ranging from 2-year degrees to PhDs).[13]

(2) Unemployed: The share of working age adults that are unemployed is defined using labor market attachment for the total working age population in the municipality. The individual

[12]This was exemplified in 2016, when a documentary series by DR (2016) showed that inhabitants in Aalborg Øst and Hasseris, two areas in Aalborg municipality were largely different measured in both income levels, education, labor market attachment, health outcomes, and their ability to make use of public health services.

[13]I use the earliest available classification of education by level from 2008 to separate educations into post- and not post secondary. Some educations change classification over time (e.g. nurses from a 2-year to 4 year professional bachelor) but it is reasonable to assume that few educations change from secondary to post-secondary and vice versa.

can be employed, unemployed, on short term leave from the labor market, or outside of the labor market (including student who do not work, early retires, chronic ill, and other groups outside the labor market).[14]

(3) Top_share: The share of individuals with income in the national top decile is found by ranking all individuals by total income[15] in the total population for each year. The municipal share is then calculated as the percentage of nationally top decile income earners compared out of the municipality population above 15. Measuring above 15 removes children from consideration who would not be expected to have incomes. A concern regarding the measure is the extend that measuring income pre-tax can skew estimates of correlations if what matters is post tax income. To the extend that taxes and deductions are monotone in the income distribution near the top of the income distribution, this will be less of a problem. A second concern with the measure is whether to take into account family sizes (e.g. by scaling incomes of families with more children/adults - see e.g. Danmarks Statistik (2017a)). Most equivalization formulas use disposable, post-tax income. As a result I cannot calculate a related measure using the available register data.

(4) Bottom_share: The share of national bottom income earners is calculated in a similar way to that of top income, replacing top decile with bottom decile income earners. This covariate similarly contains potential measurement error due to lacking tax-payment information, and no equivalization.[16]

(5) Non-western: I include the share of non-western first and second generation immigrants out of the total population as a Danish equivalent of the share of minorities in the US. Lacking proper measures of nationality in the Danish income distribution, I use measures of immigration status from Statistics Denmark.

[14] A possible improvement on this measure would be to not just measure the share of unemployed, but the share of individuals who are long-term unemployed. A problem in measuring the share of long-term unemployed with my available data is that long-term unemployed individuals might not be registered as unemployed, but shift between positions in and outside the labor force. I therefore would only be able to measure long-term unemployment with risk of substantial measurement error.

[15] Total income includes all labor earnings, personal business income, transfers, and income from wealth.

[16] Another relevant variable would be the share of the population living under a measured poverty line. This measure, however, would require equivalizations and after tax-income which have not been available for this study. I therefore leave the investigation of the prediction power of a more suitable measure of poverty to further study.

(6) Teen_mothers: The share of teen mothers to population is calculated by counting the amount of mothers who are younger than 20, and dividing with the total population. The measure relies on an indicator of the number of children in each family in the registers. A mother is indicated to have her own family if she has a child. Therefore, even if she lives with other relatives, she will be counted as a mother.[17]

(7) Teen_births: The number of teen births per 100.000 individuals is calculated using the same data as teen mothers. A teenage birth is recorded if a mother's family increases by one child while the mother is between 14 and 20. The number per municipality is scaled by population size in the municipality.

(8) Gini: I measure within municipality income inequality using the within gini coefficient for equivalized disposable (post-tax) income using data from Statistics Denmark.[18] To estimate the extend of within area inequality (which relates to the discussion of Aalborg Øst versus Hasseris), I utilize gini coefficient estimates at the municipal level collected from Statistics Denmark. The estimates are available from 1987 and I average over the period 1987-1993.[19]

(9) Crimes: The number of crimes reported to population is the sum of criminal law offences (including sexual offence, violent crime, and property crimes) by municipality. The number of crimes reported is only available from 1990. I therefore take an average over the years 1997-1993 of the number of reported crimes to approximate the distribution of criminal activity in prior periods. As crime patterns and waves changes over time the estimate may to some extend be biased. Crime is, however, a significant part of both the social disorganization and compound disadvantage theories, and so even a slightly biased measure is preferable to none.

All of the above covariates are measured at the municipality level. To investigate the hypotheses about neighborhood effects, I optimally require information at the neighborhood level of both covariates and intergenerational mobility. In chapter 3 I discuss the availability of data for the intergenerational mobility estimates. With the available data and sample sizes, I am

[17] A separate fertility register contains actual information about mother- and fatherhood. I have not have access to this register. The registered numbers may therefore deviate slightly from the actual share of teenage mothers.

[18] The gini coefficient calculates the difference in cumulative income share for each observed income rank compared to a perfectly equal distribution of income. With total population data the gini coefficient is calculated from the summation over all families. See Danmarks Statistik (2017b).

[19] The measure is fairly invariable to including more or less years, see appendix E.1.

not able to perform analyses at levels that compare to those in e.g. Wilson (1987) or Sampson (2012). Rather, I am confined to investigating effects at the municipal level. Therefore, any results that I find relate only to neighborhood effects at the aggregated municipality level. This leaves work for future research. A second concern is that segregation within municipalities (such as by immigrant status) is not measured. This however, also requires information about location at lower than municipality level in the period of interest (1980-1990). As I can only locate individuals within municipalities and not lower levels of segregation, I cannot calculate within municipality segregation (see e.g. PSC 2018). The following correlations therefore shows the relation between municipality level segregation to national shares, and not within municipality disparities. I also leave for future research the importance of local institutional resources, such as availability of local education institution, and political influence for future research.[20]

4.4 Correlations - Or, The Descriptive Findings

In the last chapter, I investigated several measures of mobility, including IGE, RRC, and transition matrix mobility measures. All measures showed variation in intergenerational mobility. In this section I focus on the correlation between the RRC estimate of mobility - the slope between parental and child income ranks, and a measure that it is based upon, which captures features of mobility: the expected rank of a child born at a given percentile rank: $E[R_c|R_p] = \hat{\alpha} + \hat{\text{fi}}_{RRC}\, R_c$, where $\hat{\alpha}$ and $\hat{\text{fi}}_{RRC}$ are the intercept and slope from rank-rank regressions at the municipal level. By inserting a given parental rank in the equation, we can calculate the expected child rank (assuming linearity of the rank-relation) for the municipality that the child is born in. This section measure deals with absolute mobility, whereas the slope coefficient, RRC, is a relative measure of mobility. A common measure of absolute mobility using this specification is expected income when born to children at the 25th income percentile, $p25$ (see e.g. Chetty et al. 2014; Chetty 2016; Rothwell and Massey 2015b; Sharkey and Torrats-Espinosa 2017).

[20] As I am writing this thesis, members of the ReDy working group at Aalborg University is mapping local availability of child care, kindergartens, schools, and higher education levels. As these data become available, they can readily be included in the analysis.

Each of the municipal-covariates captures aspects of community resources or disadvantage. Figure 4.1a shows correlations obtained from regressing standardised (mean 0, variance 1) RRC on standardized covariates with dots. Confidence intervals for each correlation is obtained from the correlation coefficient standard errors. The correlation coefficients in the plot are largely insignificant. The only variable with a substantially, statistically significant correlation is number of crimes in the local area. More reported crimes is positively related to the share RRC slope. That is, higher amounts of reported crimes in the local areas correlates negatively with relative mobility (the inverse of the RRC). The fact that none of the other variables even have correlations with the measure of mobility suggests that there is little effect of municipal level covariates on the direct relationship between parents and children.

Figure 4.1b shows correlations with absolute mobility. All covariates, except share of bottom 10 income earners and the gini coefficient appear to correlate with absolute mobility. More unemployed individuals is related to lower absolute upward mobility, as is the share of teen mothers, teen births, the share of non-western first and second generation immigrants, and the amount of crimes in the local area. The share of top income earners, as well as the share of individuals with beyond secondary education is positively related to upward mobility; given the parental income rank, children born in areas where there are more highly educated and more top income earners generally fare better than children in areas with less.

The findings in this section points in several directions. Firstly, the relative mobility estimate does not appear to relate substantially to any of the variables suggested by the neighborhood effects literature. The fact that the absolute mobility estimates does correlate with most of the neighborhood variables, suggests that there is some effect of neighborhoods on where in the income distribution children ends up. However, it does not relate directly to how parents and children's incomes relate to each other. Rather, any neighborhood effects appears to not be moderating the relationship between children and parents. It is important to note that these findings are *correlations*. In particular, correlations are linear estimates of relations. While it do not investigate it here, it is possible (and in line with the underlying theories) that particular relationships are in fact non-linear, and would not be detected from a linear estimator. I leave such exploratory investigations for future research. It is, in addition, possible that the municipal level correlations we observe are driven by self-selection - parents with lower expected child outcomes self-select into areas with higher shares of teenage births and bottom decile income earners. Similarly parents with children having higher expected income ranks may self-select into areas with lower levels of teen births, etc. Using only the correlations we cannot discern correlation from causation. In the next section, I follow this line of inquiry, and develop a model for estimating whether the municipality a child lives in fact has causal effects on children's outcomes.

4.5 Estimating Neighborhood Effects

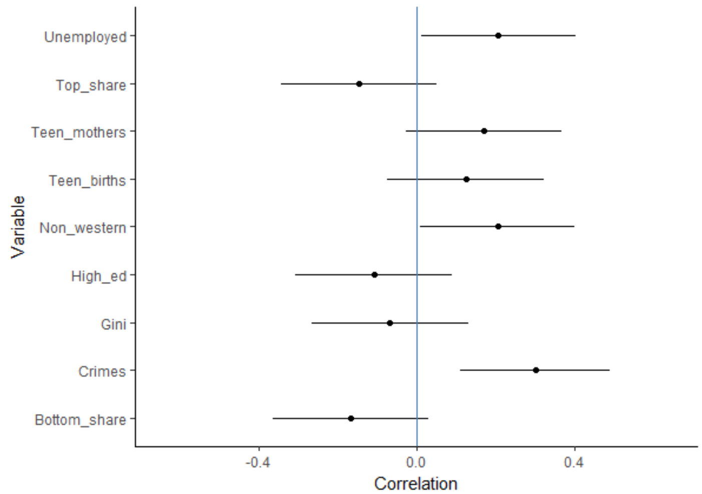
The literature review has highlighted two important elements: Firstly, that there are theoretical reasons to believe that in particular bad neighborhoods can have negative influences on individual mobility. Secondly, that empirical investigations of neighborhood effects in relation to long-term outcomes generally have been suggestive of some effects of neighborhoods, although the most reliable studies that use experimental evidence are from the US, and primarily investigate highly disadvantaged parts of the population. Before turning to estimation of neighborhood effects in this study, I will first expand the model proposed in this section to give an account of how we can interpret neighborhood effects in relation to the original model.

The original equation from the theoretical framework above was

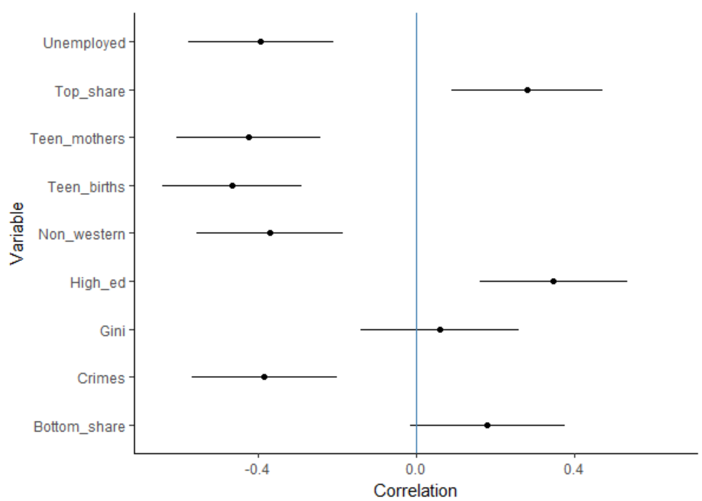
$$Y_c = \alpha + \beta Y_C + \epsilon. \tag{4.1}$$

In the error-term, ϵ, are all of the effects that may covary with individual characteristics, and on which the estimate of β will depend. Making a strict assumption, we may argue that neighborhoods have an effect per year spend in the neighborhood on the child, which is separable from that of parental income, such that $\epsilon = N_K + \nu$, where ν is another error-term, and N_K is the cumulative effect of exposure from living in the neighborhood K through

Figure 4.1: Correlations between municipal rank-rank correlation/expected child outcome (parental rank 25) and covariates.



(a) Rank-rank correlations



(b) p25 correlations

Note: The figure shows correlations between municipality RRC and p25 estimates and covariates with dots, and 95 percent confidence intervals indicated with lines. Correlations are calculated by regressing standardized RRC and p25 (mean 0, variance 1) on similarly standardized covariates. As both variables in the regression have the same standard deviation, the coefficient from the regression is the correlation (see the derivation in chapter 3), and standard errors of the correlations are the regression standard errors. 95 percent confidence intervals are obtained under assumption of asymptotic normality of correlation coefficients as $\rho \pm 1.96\sigma$, where σ is the standard error.

childhood.

$$Y_c = \alpha + \beta Y_C + N_K + \nu. \tag{4.2}$$

Splitting up the term N_K in year-wise effects, we get $N_K = \sum_{a=1}^{A} D_{a,K}$, where a is the present age, and A is the latest year in which the neighborhood can have an effect. Assume now that a family moves at age g to destination neighborhood d. Within each year the child spends in another neighborhood, we assume that the effect of being in that neighborhood is $D_{a,d}$. If there are no confounders, then the neighborhood effect on the child will be $N_{K,d} = \sum_{a=1}^{g-1} D_{a,k} + \sum_{a=g}^{A} D_{a,d}$. Writing the effect from the destination neighborhood as the difference from the original neighborhood (adding and subtracting effects of the original neighborhood), we get $N_{K,d} = \sum_{a=1}^{A} D_{a,k} + \sum_{a=g}^{A} D_{a,d} - D_{a,k} = N_K + N_d$. N_d is now the effect of differing neighborhood quality. Inserting in the original equation,

$$Y_c = \alpha + \beta Y_C + N_K + N_d + \nu. \tag{4.3}$$

we get an equation that contains the effects of living in differing neighborhoods, assuming that N_d and ν independent. Since we are interested in intergenerational mobility, the relevant feature of difference between neighborhoods would be the differences in expected outcome (rank) of the children and given parental income in either municipality. Letting the expected outcome for a child with parental income at rank p in origin (o) and destination (d) neighborhoods be $\bar{y}_{op}$ and $\bar{y}_{dp}$, we can rewrite to $N_d = \sum_{a=g}^{A} \gamma_a \bar{y}_{dp} - \bar{y}_{op} = \sum_{a=g}^{A} \gamma_a \Delta_{odp}$. Here γ_a is the expected effect of living in neighborhood from age a.

This regression is of course only interesting for comparing movers (the difference terms will be zero for all non-movers). Naturally, the neighborhood that the family moves to need not be independent of the error term in the model ($cov(\Delta_{odp}, \nu) \neq 0$). In this case, an OLS estimator of γ_a would be biased by the direction of covariance with omitted variables. NIf we could control for all relevant variables (with correct functional form) we would have a saturated model, and we could estimate the causal effect of neighborhoods (Angrist and Pischke, 2009, p. 48-59). However, it is likely that some family characteristics (we must assume the characteristics to be at the family level, as we are specifying neighborhood effects) are unmeasurable and co-varies with neighborhood quality. Examples can be high achieving parents choosing to move to areas with better schooling options and non-curricular activities, which would inflate the estimate of the effect of the neighborhood difference, if this also leads to higher expected mobility in general.

Chetty (2016), using a similar approach to investigating neighborhood effects, provides an assumption that can separate causal effects from selection effects. First, assume that the observed neighborhood effect at age a is additive in true neighborhood effect (μ_a) and the effect of selection (s_a), $\gamma_a = \mu_a + s_a$. If it is the case the that selection effects do not vary with age $s_a = s$, then the effect of spending one more year (from $a+1$ to a) in the destination neighborhood can be found as $\omega_a = \gamma_a - \gamma_{a+1} = \mu_a + s - \mu_{a+1} - s = \mu_a - \mu_{a+1}$. This provides us with an opportunity to estimate the causal effect of neighborhoods through the effect of staying one additional year in each neighborhood.

Estimating an equation (for movers) as above is confined to tell us something about the individuals that actually do move. It cannot tell us anything about those that do not. The assumption that there are no variables that may confound the effects, can be problematic. Assume for example that the reason why a family moves is that both parents find a new job and higher paying job, and thereby increases their investment in the child - an effect that is internal to the family, not an outcome of network relations, better available resources, nor better networks or local community social control (some of the theoretical effects from the neighborhood effects literature). In this case, we have a confounding variable that is not accounted for by permanent parental income. Similarly, changes in parental characteristics, such as marital or relational status, can as well effect the long run expected outcome of the children. We can, however, check if such changes have an effect, by including any changes from before to after a relocation by including such characteristics as covariates in a regression. Another potential problem is that families differ to a substantial extent, which cannot be accounted for by observable variables, in a way that otherwise correlates with the movement. To account for such changes, we can include family fixed effects in the relevant regression specification, relying on differences in outcomes by age of children within families. This, however, requires a substantial amount of included covariates in the regression, suggesting that estimation problems may be severe.

One problem with the estimation of the above effects is that we cannot discern if the theoretical explanations for neighborhood effects actually are important. It is, however, potentially possible to discern whether particular theories have more merit than others. In particular, if we estimate separate coefficients for individuals moving to significantly better neighborhoods (more than one standard deviation difference in expected outcome at the parental income percentile), the relative deprivation theory would posit that the children from worse neighborhoods may not be able to access all of the resources in their new neighborhood, and therefore the effect might be smaller for this group than those at higher levels. In a similar way, if a child from a better neighborhood moves to a worse one, the compound disadvantage theory would posit that the effect would be less than for the average kid. Therefore we might expect less conversion (a smaller coefficient) for children moving to areas significantly worse than their present in terms of expected outcomes (1 standard deviation or more in expected outcome at parental income difference). Opposite this finding, institutional and social isolation theories would make us expect relatively stronger effects for movements to worse areas.

Before attempting to control for the various specifications above, we must note that these controls are only relevant to the extent that the initial tests (first regressions) show the expected pattern; a steadily declining effect of moving to better neighborhoods as the child grows older. This is the baseline test I employ to test for the existence of neighborhood effects on mobility.

4.6 The Effects of Neighborhoods

The main specification that I use to identify neighborhood effects is

$$R_c = \alpha + \beta_{RRC} R_p + \sum_{k=1}^{K} N_k + \sum_{m=8}^{A} b_m \Delta_{odp} + \omega \tag{4.4}$$

The second term estimates the parental income correlation from before. The third term is the set of origin indicators, one for each municipality. An indicator takes the value 1 if the child has origin in that municipality before moving. The fourth term is the one of primary interest. The term is comprised of an indicator for moving at age m, and the difference in expected outcomes at the child's parental income rank, p.

Table 4.1: Summary statistics for the samples of children who move once before they are 15 or 20 and their parents.

Movers by age 15 ($n = 23,074$)

Statistic	Mean	St. Dev.	Pctl(25)	Pctl(75)
Age at move	11.127	2.276	9	13
Child income rank	0.483	0.297	0.215	0.742
Parental income rank	0.522	0.296	0.261	0.785
Predicted outcome - origin	0.503	0.079	0.438	0.569
Predicted outcome - destination	0.507	0.077	0.443	0.571
Difference	0.004	0.026	−0.012	0.020

Movers by age 20 ($n = 35,120$)

Statistic	Mean	St. Dev.	Pctl(25)	Pctl(75)
Age at move	15.746	4.424	11	20
Child income rank	0.505	0.295	0.245	0.766
Parental income rank	0.531	0.289	0.280	0.786
Predicted outcome - origin	0.508	0.076	0.447	0.571
Predicted outcome - destination	0.505	0.075	0.445	0.568
Difference	−0.002	0.027	−0.019	0.015

I consider two different samples of moving children. The first is children who move once between the ages of 8 and 15, the second children moving once between 8 and 20. The lower limit (age 8) on when children's moves are measured follow from the available data. The earliest possible year that the cohort of 1973 can be observed is 1980, at the age of 7. This means that the earliest year a child in this cohort can be observed in the data to have moved is at age 8. A theoretical concern about the earliest age of movement relates to the early childhood development literature (Cunha and Heckman, 2007; Heckman, 2008b); studies in this literature contents that there are substantially higher returns investing in children in their early years (often mentioned as up to age 5 or 6). As a result, the assumption of constant selection effects might be violated for earlier ages, in particular if parents are more likely to move around in order to provide the best environment for the child in its earliest years. For the first sample, these experience moving during the period in which I originally measured parental income (less the first year in which the child is 7). The second, and larger set of movers, comprises children moving throughout their teenage years, most likely in connection with their parents. I choose to use two samples as the first sample is somewhat restrictive on when the child moves. The second, on the other hand, may also contain children that move independently of their parents, and so are actively choosing their own neighborhood. For the early sample it is more likely that the argument of fixed selection effects apply.

Table 4.1 contains summary statistics for each sample. We see that the number of observations differ between the two groups. In the first group the sample consists of 23,74 children, and in the second the sample consists of 35,120 children. Movers in the early movers group is, as expected, lower at 11.1 years versus 15.8 years on average. The early movers in addition also tend to do slightly worse than the later movers, with average outcome ranks of 48.3 against 50.5. The early movers, however, tend to move to areas that are slightly better in terms of expected child rank given parental income rank; the difference between municipalities is .4 ranks for the early movers, and -.2 ranks for the later ranks.

I next perform the regressions for each set of movers as specified in equation 4.4. The regression is performed by OLS with heteroskedasticity robust standard errors (Angrist and Pischke, 2009, pp. 294-308). The results of the regressions is found in table 4.2. Firstly, we see that in spite of inclusion of origin fixed effects, the parental rank estimate is rather similar to that we originally estimated for the national level. This suggests that while any municipality effects may exist, they do not interfere substantially with the estimated linkage between parental and child income ranks. More interesting, however, is the estimated coefficients on the age of movement. I have argued that with origin fixed effects, we should expect steadily declining age of movement effects. Looking first at the sample of individuals that move once between the ages of 8 and 15, we find a relatively stable movement coefficient around 0.4 for moving from age 8-10 (all statistically significant). By age 12 this pattern breaks, and the effect drops substantially to .251, followed by near zero estimates at age 13 and 14. Neither effect is statistically significantly different from zero. Finally, at age 15 we see a large negative effect of movement of -0.771 (which is statistically significantly different). We see similarly erratic movements in effects of moving from the second regressions, where effects of movement swings from positive to negative from the age of 12 until age 20 (with only effects

Table 4.2: OLS regression of child income rank on parental rank and age-at-move interactions with origin-destination municipality differences in expected outcome

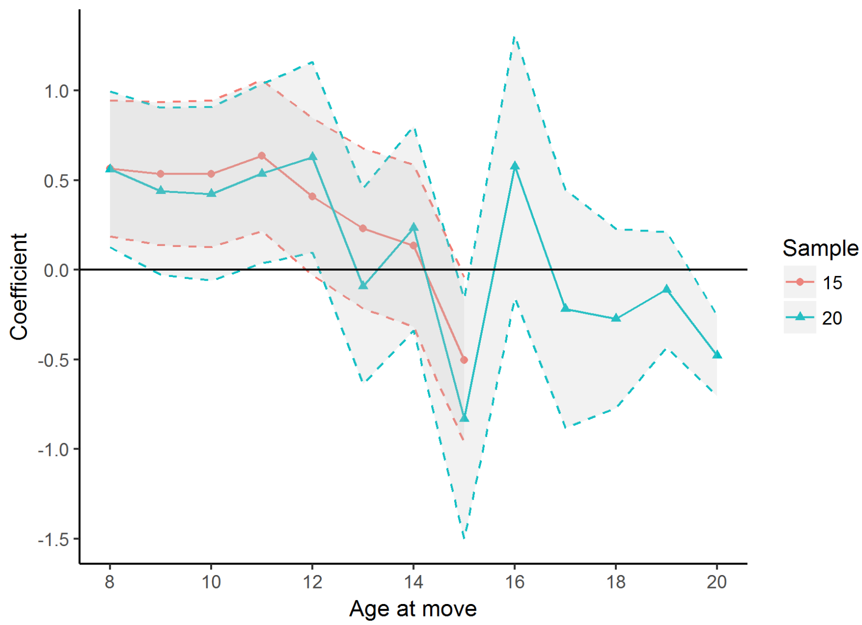
	Dependent variable:	
	Child Income Rank	
	Movers by age 15	Movers by age 20
	(1)	(2)
Parental income rank	0.287***	0.257***
	(0.006)	(0.005)
8:diff	0.431**	0.448**
	(0.188)	(0.218)
9:diff	0.403**	0.306
	(0.199)	(0.236)
10:diff	0.399**	0.296
	(0.202)	(0.243)
11:diff	0.452**	0.413
	(0.210)	(0.251)
12:diff	0.251	0.538**
	(0.217)	(0.265)
13:diff	0.064	-0.186
	(0.223)	(0.275)
14:diff	-0.089	0.045
	(0.224)	(0.287)
15:diff	−0.771***	−0.973***
	(0.230)	(0.340)
16:diff		0.414
		(0.373)
17:diff		-0.363
		(0.338)
18:diff		-0.370
		(0.253)
19:diff		0.004
		(0.161)
20:diff		−0.775***
		(0.105)
Constant	0.304***	0.329***
	(0.017)	(0.015)
Origin fixed effects	Yes	Yes
Observations	23,074	35,120
Adjusted R^2	0.087	0.068

*p<0.1; **p<0.05; ***p<0.01

Note: The table contains regression estimates for a sample of children who move once from the age of 8 to 15, and (2) from the age of 8 to 20. Child income rank (in the national income distribution in the late 30s) on parental income rank (among other parents of children in the sample at the national level, measured when the child is between 7 and 15). The regression also includes a covariate of the difference (*diff*) in expected child income rank between the origin municipality (where the child moves from) and the destination municipality (where the child arrives). The covariate is interacted with age dummies for moving between at each age 8-15 (1) or 8-20 (2). Both regressions also include origin municipality fixed effects. The regression is estimated with heteroskedasticity robust standard errors.

of movement at age 12, 15, and 20 statistically significant). Figure 4.2 shows the coefficients with 95 percent confidence intervals for the 15 and 20 groups. The plot underscores the story, with erratic movements in coefficients by age and marginally or non-significant coefficients at most age-levels.

Figure 4.2: The effect of moving to a better municipality by child age at move



Note: The plot shows the coefficients on origin-destination municipality expected child outcome by age of movement. The coefficients are obtained from regression on two sample of children and parents which move (1) when the child is between 8 and 15, and (2) when the child is between 8 and 20. The regression specifies child income rank as a function of parental income rank, a set of origin-municipality fixed effects, and the set of interactions between age at move with the difference in expected child income rank in the origina and destination muncipality. Gray ribbons indicate 95 percent confidence intervals around the estimated effect based on heteroskedasticity robust standard errors.

The expectation for the neighborhood effects regression is that we would be able to see that children moving to better municipalities converge towards the expected income levels of the new municipalities. This would show up with a positive coefficient on the expected outcome difference, declining in age. However, the non-statistical and erratic coefficients in the sample of movers suggests that there is no such conforming. Another way to investigate if there is signs of such convergence is non-parametric estimation of the relation between child income rank and the expected outcome difference (Chetty 2016 uses a similar approach to show that there actually is convergence in the US).

To make the non-parametric investigation, I firstly redefine the neighborhood effects model in equation 4.4. The goal is to define a function based on indicator variables, and the difference variable, such that we can express the expected child outcome ranks conditional on the origin-destination difference. By defining a function with indicator variables, we can condition on the parental income level and origin neighborhood simply by calculating the differences from parental rank-origin group means. Let the indicator variable for being in parental group p, and from origin o be ω_{po}. The relationship between child outcome rank and the difference in expected outcomes from moving at age m is then defined by

$$R_c = \omega_{po} + b_m \Delta_{odp} + v \tag{4.5}$$

The conditional relation between child outcome rank and the difference (at age m) can then be found by regressing each variable (R_c, Δ_{odp}) on the set of indicator variables ω_{po} and regressing the two residuals from the regressions.[21] We can then plot binned scatterplots of the two sets of residuals to investigate, non-parametrically, if there are signs of convergence (linear relationships declining with age group).

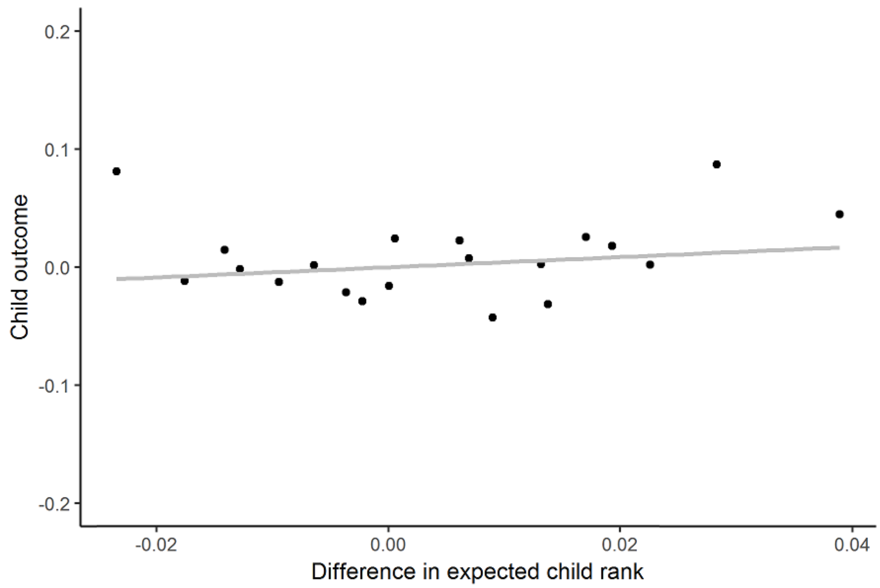
Any regression on a set of indicator variables corresponds to demeaning within the indicator variable groups. I therefore find the two sets of residuals by demeaning within parental rank-origin groups. As parental ranks is a continuous measure, I choose to group the parental income ranks in deciles. The result is a set of $10 * 98 = 980$ indicator variables. Using percentiles instead would give a total of $100 * 98 = 9,800$ indicator variables, close to the number of observations, making the demeaning nonsensical.

Figure 4.3 shows the relation between child outcome rank and difference in origin-destination expected outcomes, conditional on parental income decile and original municipality, for movers at age 10. Each dot indicates the mean outcome within a ventile of difference in expected outcome. Under the assumption of adaptation to the destination municipality, we should see a tendency for the mean values to have a positive slope. What we actually observe in figure 4.3 is what appears to be a random pattern in expected value among the children who move when they are 10. This pattern is exemplary for movers at all age groups. Figure E.2 in

[21] This is a result of the Frisch-Waugh-Lowell theorem (Frisch and Waugh, 1933).

appendix E.2 shows similar binned scatterplots for movers at age 8-15 for the individuals who move before turning 16.

Figure 4.3: Binned scatterplots of child income-rank by expected origin-outcome difference ventiles, controlling for origin municipality and family fixed effects. Group of movers at age 10.



Note: The plot shows binned scatterplot values. For each ventile of difference in origin-destination expected child rank, the average outcome of children in the age-group has been averaged. The grey line shows indicates a linear approximation to the points in the figure.

In this section we have seen that there are no strong indications of neighborhood effects at the municipal level. In fact, it appears that relations between children's outcome ranks and differences in origin-destination quality is random. This stands in contrast to the findings by Chetty (2016) who find that *each year* spent in a 1 percentile better neighborhood (difference in expected outcome rank) increases the outcome of the child by 0.04 percentiles. These findings suggests that any effects of neighborhoods on intergenerational mobility is not to be found at municipal level.

Chapter 5

Discussions

In this chapter, I relate the findings of each of the three main chapters of the thesis to the original research questions. I then discuss particular limitations of the thesis, as well as the future research avenues that these and the findings might suggest. The main research question of the thesis is

> Is equality of opportunity affected by neighborhoods in Denmark?

In order to answer it, I developed three sub-questions:

1. What characterizes the spatial distribution of opportunity in Denmark?
2. Why might particular groups defined by geography experience higher or lower degrees of equality of opportunity?
   - What role do neighbourhood characteristics play as predictors of equality of opportunity?
3. What are the effects of changing neighbourhood for equality of opportunity of families?

Answering each question requires first defining what equality of opportunity entails. In chapter 2 I introduced three modern notions of equality of opportunity: Formal equality of opportunity (Arneson, 2018), Rawlsian fair equality of opportunity (Rawls, 1999), and Roemerian algorithmic equality of opportunity (Roemer, 1998). When introducing them, I showed that none of these notions directly relate to the typical measures of intergenerational mobility. In fact, studies estimating mobility measures are under-defined with respect to equality of opportunity if they cannot determine relevant policies and circumstances. The problem can be viewed as one of measurement, or one improperly applied philosophic notions. The source of the problem is irrelevant here. The distance between philosophy and economics is smaller for neighborhood effects. Under Romerian equality of opportunity it is sufficient to show that neighborhoods causally affect intergenerational mobility, to show that equality of opportunity is affected by neighborhoods.

With the goal of investigating whether neighborhoods have causal effects of neighborhoods, I set out in chapter three with measuring observed mobility by neighborhoods. In the chapter, I discuss various characteristics of the intergenerational mobility literature, including the substantial concerns regarding measurements of mobility. These concerns lead me to the first major limitation of the study, the definition of neighborhoods. Measuring mobility requires data support for performing regressions, or estimating probabilities of transition matrices. The data that I employ in this thesis is a subset of Danish register data, which has been made available from Statistics Denmark through the IKE research group at Aalborg University. The data spans the total Danish population in the period 1980-2012 and allows for linking parents and children, as well as measuring their total income over longer periods of time. Unfortunately the data does not allow for measuring the location of families at spatial aggregations lower than municipalities. The definition of what constitutes neighborhood varies with application. Wilson (1987) and Sampson (2012), two of the neighborhood effects scholars, work with neighborhoods at block, or defined by inhabitants beliefs of borders between areas. Chetty et al. (2014) and related papers define neighborhoods e.g. as commuting zones in the US. Regardless of what is a reasonable measure, it is clear that municipalities are an administrative unit, which may not exhibit the characteristics that are relevant for neighborhood effects studies on intergenerational mobility. This provides the first avenue for future extensions of the research in this thesis: investigating the extent of mobility differences across various other spatial aggregation levels.

I find that mobility does in fact vary with location. While this is not sufficient to argue that there are in fact differences in opportunity across neighborhoods defined as municipalities, it is indicative, and rules out the zero-case of no spatial variation in mobility and equality of opportunity.

In chapter four I proceed to answering sub-question two. The literature on neighborhood effects (summarized e.g. by Small and Newman 2001 and Wodtke et al. 2016) provides several theoretical explanations of why neighborhoods might matter for equality of opportunity in Denmark. These include (but are not limited to) social isolation theory, social disorganization theory, institutional resources, compund disadvantage theory, and relative deprivation theory. Common among the theories is that most emphasize highly deprived areas, and often in inner city areas (a point noted by Small and Newman 2001). This casts doubt on their applicability at the level of Danish municipalities. It is not unlikely that new theoretical models may need to be developed for the significantly different social context that is often apply in Denmark, a country characterized often by high levels of social trust and government support for the least advantaged.[1] Regardless, the US-centric neighborhood theories, and the supportive causal inference evidence e.g. from the Moving To Opportunity Experiment, as well as the work spurred by Chetty et al. (2014) (e.g. Sharkey and Torrats-Espinosa (2017)) suggests that neighborhoods may in fact affect intergenerational mobility, and thereby equality of opportunity.

Main contenders among the theoretical factors that could explain effects are the concentration of local disadvantage, the availability of institutional resources, as well as local social cultural norms enacted partly through available mentors, peers, and members of the community. The relative deprivation theory suggests that higher inequalities in the local area can attenuate less advantaged children's lives. In the analysis that answers the second part of the second sub-question, I include a set of 9 variables, each designed to approximate factors from the various theories; The share of highly educated individuals in the municipality, the share of unemployed, the share of non-western first and second generation immigrants (to

[1] The infamous reference on government expenditures and social welfare capitalism characteristics is Esping-Andersen 1990. The Danish welfare state has seen significant changes since 1990, but it is worth noticing that the children in this study actually grew up in the period that Esping-Andersen emphasized in his book.

approximate the US equivalent: racial segregation), shares of teen mothers and teen births, the gini coefficient measuring income inequality, and finally the number of reported crimes in the local area. These variables touch on aspects of the social disadvantage theory and relative deprivation theory in particular. Unfortunately, many limitations were present when searching for relevant variables at the municipal level.

The primary limiting factor is that several variables are not easily measurable or were not available at the time of this study. An important factor in social disorganization theory is the community social control, which is an enactment of social values in relation to particularly children. Measuring these enactments requires knowledge about social networks - who gossips with whom over the hedges - and what are the cultural values in the local society. Measuring social networks has seen an increased interest in the social sciences as new simulation theories become more powerful and can be linked with online social data, such as interactions on twitter and other social media (Andris, 2014). Social values are an important aspect of the ethnographic parts of the neighborhood effects literature (Sampson, 2012, chapter 2), but scaling is difficult. As noted by (Oreopoulos, 2003, p. 1537)."*Not everyone in a deprived neighborhood is a gang member.*" and, we might add, not everybody engages with the gang members. Scaling to municipalities may not be possible with this type of research, and what is needed, is instead information, for example about peer interactions among children, and the effect of bringing new information into the closed loop. This approach goes hand in hand with analytic research by e.g. Akerlof and Holden (2016) on identity economics (Akerlof and Kranton, 2000), which suggests that individuals (and also children) respond by adapting identities to the information that the acts of others confers to them. Empirical research in this direction already appears in the form e.g. of peer effect studies, but the intersection with equality of opportunity research may be fruitful. Another strand along these lines is that local resources theories. Relevant factors may be the share of local high quality child care entities, higher education institutions, the opportunities for local political engagement. In the face of limited data, I have not included these aspects here. But, researchers at the Center for Regional Dynamics and Disparities are presently compiling information about such institutional resources. Naturally, it would be interesting to extend the analysis of the thesis with this information.

Many other covariates could be included according to the neighborhood theories. While these have not been included in the correlational analysis of chapter four, it remains the case that the variables that were included did in fact correlate with absolute intergenerational mobility, the expected income rank of children born at the 25th parental income rank. This was not the case for relative mobility - the statistical relation between parental and child outcomes. By nature, the correlations are descriptive statistics, and they cannot tell us whether the observed relations are true, or driven for example by families self-selecting into specific areas. The answer to the second part of sub-question two is that measuring with a limited set of neighborhood effects related variables, these variables do predict absolute mobility differences between local areas. As for the initial intergenerational mobility estimates, we cannot infer from these correlations that equality of opportunity is similarly predicted by neighborhood covariates, but is is indicative.

The final sub-question takes on the question of causal effects. In chapter four, I develop a caual inference framework based on the work by Chetty (2016): Among a sample of families that move once, I focus on the difference in expected outcomes of the child, had it grown up in either municipality. A municipality is better, if the child would have higher expected income rank there. Among the sample of movers, it is likely that the families change area for a reason - that is, they choose the area that they move to. Under the assumption of constant selection effects, however, it is possible to estimate the causal effect of moving one year earlier to the local area, and separate this yearly additional effect from the selection effect. This is the framework I use to investigate neighborhood effects at the municipal level, using two samples of movers among the original families in the register data: families that move while the child is between 8 and 15, and families that move once while the child is between 8 and 20. The main expectation under the assumption of neighborhood effects is that for each additional year spent in a neighborhood, the more the outcome of the child (on average) will adapt to that of the destination municipality. Estimating the model, I find that the expected pattern of smooth decline with age of movement is not present in the data. Rather, at the municipal level, I find an erratic pattern over the ages of 8-20 (and 8-15). The test of neighborhood effects is rejected at the municipal level, and I can answer the third sub-question in the negative. The result of the findings in the second part of chapter four provides the final conclusion to the main research question: Equality of opportunity is not affected by municipalities in Denmark.

Naturally, as is suggested by limitations in the former sections, this is not a final rejection of any neighborhood effects on equality of opportunity in Denmark. The first reason is that the causal analysis has been performed at the municipal level. The neighborhood theories apply to lower levels of spatial aggregation, indicating that there may be still be effects that are obscured in the municipal level analysis. Secondly, having rejected the hypothesis regarding adaption in child outcomes, we are left with additional data that we can make us of exploratively (Tokey, 1980). From the analysis, the predicted adaptations appear to follow a pattern that, for example, could be consitent with crucial life cycle points for children (see e.g. Currie 2001). Under this assumption, children do not experience a constant cost of movement. The cost rather varies with age, and a single cost-point in time may be severe, such as for example near the end of primary school (consistent with the negative effect at age 15). Under this assumption it is possible that neighborhood effects do exist, but we cannot use the framework I deviced to find them as it is obscured by costs of movement that are a function of age itself. This is an ex post rationalization on already observed outcomes. It cannot help us determine if neighborhood effects exist, but are hidden by other move related functions of age.[2]

Returning to the findings of the thesis, taking them at face value, we have found the following: neighborhoods do not have an effect on intergenerational mobility. From the discussion of equality of opportunity it is clear that this corresponds to the statement that neighborhoods do not affect equality of opportunity in Denmark.[3] This naturally begs the question of *why not*? Chetty (2016); Chetty and Hendren (2016) find that neighborhoods (measured at the commuting zone and county level) do in fact have effects on mobility in the US. This finding corresponds with the general finding from the intergenerational mobility literature that the

[2]One way to test if this ex post rationalization is in fact a likely explanation is to redo the analysis for another sample. One way of doing this, would be to collect comparable data for a sample of children that have grown up 10 years later, and reestimating the model. Observing a similar pattern for this second sample, we may increase our faith in the theory. A second problem then arises, however. Unless we can specify the second age-dependent function this means that we cannot estimate the causal effect at movement.

[3]And by tangent, assuming that the municipal level results are generalizable, that any differences in observed outcomes between areas are simply a function of self-selection.

US experiences markedly lower rates of mobility than Scandinavian countries. The epitome of this discussion is the “Great Gatsby Curve,” (figure A.1 in appendix A.1) which relates a country’s estimated mobility to its level of income inequality - higher income inequality is related to less mobility. Returning to the initial discussion of equality of opportunity it is clear that one aspect that can lead to differences in equality of opportunity is public policy: societies that are better at equalizing across relevant types (constellation of circumstances) will have more equal distributions of opportunity. The social capitalism literature (Esping-Andersen, 1990) and the Varieties of Capitalism (Hall and Soskice, 2001) literatures provide elaborate frameworks from which we can attempt to understand institutional differences between countries. This, of course is yet another avenue for future research. An attempt to investigate whether institutional characteristics in fact are a causal driver of eliminating or maintaining neighborhood effects is to exploit the (arguably exogenous) variation in public policy to estimate effects on intergenerational mobility and neighborhood effects, and then in turn the effects on equality of opportunity.

In this section I have outlined the findings of the thesis. I have found that while I could not directly infer whether distributions of opportunity varied with neighborhoods, the observed differences in mobility are indicative of some relation between municipalities and equality of opportunity. Predictive power of variables measured at the municipal level, and drawn from neighborhood effects studies, further supports the suggestive evidence. A test of causal effects of neighborhoods at the municipal level, however, found no evidence in favor of causal effects and was rejected. The result of this is that I find that equality of opportunity is not affected by neighborhoods in Denmark, conditional on the characteristics of research design. I have indicated various potential confounders in the research design, stemming partly from data limitations, and partly from operationalization. These potential confounders, as well as the findings of the study gives rise to a range of future avenues of research.

Chapter 6

Conclusion

Equality of opportunity is a fundamental concept for many modern conceptions of justice, and a topic of intensive research. In this thesis, I answer the research question "*Is equality of opportunity affected by neighborhoods in Denmark?*" through a sequence of three sub-questions: (1) What characterizes the spatial distribution of opportunity in Denmark? (2) Why might particular groups defined by geography experience higher or lower degrees of equality of opportunity? What role do neighbourhood characteristics play as predictors of equality of opportunity? (3) What are the effects of changing neighbourhood for equality of opportunity of families?

In order to answer the reserach questions, I initially investigate three philosophical notions of equality of opportunity: formal, fair, and algorithmic equality of opportunity. I show that in order to answer the research question affirmatively, it is necessary to document that neighborhoods causally affect intergenerational mobility across municipalities. Documenting spatial differences in measured levels of intergenerational mobility is not sufficient to show that equality of opportunity is affected by neighborhoods.

I utulize Danish register data in order to answer the empirical first question, drawing on the theoretical and methodological background of the intergenerational mobility literature. The data spans the period of 1980-2012, allowing me to link children and parents, as well as measuring total income of parents when the child is 7-15, and the child's income in its late 30's.

Using a variety of mobility measures I document substantial variation in intergenerational mobility across Danish municipalities, the smallest geographic entity that I can link individuals to. The finding of variation in mobility is not sufficient to argue that the spatial distribution in opportunity similarly varies. However, it opens up the possibility that neighborhoods can affect equality of opportunity. I next discuss the theoretical foundations from which we can argue that neighborhoods might affect equality of opportunity, drawing on the neighborhood effects literature. Several theories, including social disorganization theory, local resource theory, and relative deprivation theories, as well as social isolation theory, provide frameworks that have been developed in the US, primarily to explain inner city deprivation. I develop a set of nine indicators that are expected to relate to effects of neighborhoods. Among them are the number of reported crimes to population, the share of highly educated individuals, share of bottom income earners, and the amount of teenage births to youth population. As for the income data, I base the majority on the available register data, measuring averages per municipality from 1980-1990. I am not able to operationalize several elements in the theories, including the complex network and social control mechanisms, as well as the availability of local resources, which are integral parts of the theories. This is therefore left for future research. I am, however, able to document that absolute intergenerational mobility at the municipal level correlates at a statistically significant level with the neighborhood effects indicators. Once again, this cannot show that equality of opportunity is affected by neighborhoods. It is possible that the correlations are driven by self-selection by families into specific areas. The correlations, however, suggest that the theoretical arguments for why neighborhoods might affect equality of opportunity in smaller geographic entities and/or when taking into account additional indicators.

The final part of the thesis investigates whether there is in fact a causal impact of neighborhoods, operationalised as municipalities, on intergenerational mobility. I develop a causal inference framework in which I use information about two sets of moving families, the first is families that move once while the child is between 8 and 15, the second families that move once while the child is between 8 and 20. The identification strategy relies on the assumption that selection effects among the movers is constant over the ages of 8-15, and 8-20 respectively. If this is the case, then the causal impact of moving to a better neighborhood can be measured as the adaptation in child income rank to that expected in the destination neighborhood, given the origin neighborhood. I expect that this adaptation is smoothly declining in the age of the child - the fewer years spent in the destination the less effect. Estimating the devised model, I find no support for the hypothesis. The coefficients on moving to different neighborhoods at different ages move erratically with age. As a result, I must reject that I find any effect of neighborhoods on intergenerational mobility in the causal inference framework.

The fact that I can reject a causal effect of neighborhoods at the municipal level also provides the answer for the main research question. Equality of opportunity are not affected by neighborhoods in Denmark. The result is conditional on the research design in the thesis. A particular concern is that municipalities are not an appropriate operationalization of neighborhoods. Future research utilizing other datasets may be able to find neighborhood effects with alternate research designs. A second concern is that children may experience costs of moving that is not constant across age of moving. If this is the case, then the causal framework cannot identify effects of neighborhoods. This also leaves room for future research. Taking the finding in the study at face value, it suggests that there are substantial differences between the Scandinavian country of Denmark, and the United States. Chetty (2016) has shown that neighborhoods do causally affect intergenerational mobility in the US. A third line of future research is to investigate further whether e.g. welfare policies or labor market institutions are driving this result.

Bibliography

Acemoglu, D. and Autor, D. (2009). Lectures in labor economics. *Unpublished manuscript, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA. Accessed on March 15, 2018 - URL: http://economics. mit. edu/files/4689.*

Akerlof, G. A. and Kranton, R. E. (2000). Economics of identity. *Quarterly Journal of Economics*, 115(3):715–753.

Akerlof, R. and Holden, R. (2016). Movers and shakers. *Quarterly Journal of Economics*, 131(4):1849–1874.

Alesina, A., Stantcheva, S., and Teso, E. (2017). Intergenerational Mobility and Preferences for Redistribution. *NBER Working Paper Series, 23027.*

Andris, C. (2014). Relevant Literature for Examining Social Networks in Geographic Space. *SFI Working Paper, 2014-04-010, Santa Fe Institute.*

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Arneson, R. (2015). *Equality of Opportunity.* The Stanford Encyclopedia of Philosophy (Summer 2015 edition).

Arneson, R. (2018). Four Conceptions of Equal Opportunity. *The Economic Journal*, pages 1–22.

Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, 49(1):3–71.

Beck-Nilsson (2018). *Bred alliance vil give småbørn lige muligheder for et godt liv.* Altinget.dk, Altinget.dk, Online article, Accessed on March 3, 2018 - URL: https://www.altinget.dk/artikel/bred-alliance-vil-give-smaaboern-lige-muligheder-for-et-godt-liv.

Becker, G. S. (1993). *Human Capital - A Theoretical and Empirical Analysis With Special Reference to Education.* Chicago University Press, Chicago, 3 edition.

Becker, G. S. and Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, 87(6):1153–1189.

Becker, G. S. and Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4(3, Part 2):S1–S39.

Betts, J. R. and Roemer, J. E. (1998). Equalizing Educational Opportunity Through Educational Finance Reform. *UC Davis Working Paper 99-08.*

Bingley, P., Corak, M., and Westergård-nielsen, N. (2011). The Intergenerational Transmission of Employers in Canada and Denmark. *IZA Discussion Paper Series 5593.*

Björklund, A. (2016). Intergenerational mobility , intergenerational effects , the role of family background , and equality of opportunity : a comparison of four approaches. *Presentation, OECD, url: https://www.oecd.org/employment/emp/OECD-ELS-Seminar-Björklund-IntergenMobilityComparison.pdf.*

Björklund, A. and Jäntti, M. (2012). How important is family background for labor-economic outcomes? *Labour Economics*, 19(4):465–474.

Björklund, A., Jäntti, M., and Lindquist, M. J. (2009). Family background and income during the rise of the welfare state: Brother correlations in income for Swedish men born 1932-1968. *Journal of Public Economics*, 93(5-6):671–680.

Björklund, A., Jäntti, M., and Nybom, M. (2017). The Contribution of Early-Life Versus Labour Market Factors to Intergenerational Income Persistence: A Comparison of the UK and Sweden. *The Economic Journal*, 127(2013):71–94.

Björklund, A., Jäntti, M., and Roemer, J. E. (2012). Equality of opportunity and the distribution of long-run income in Sweden. *Social Choice and Welfare*, 39(2/3):675–696.

Black, S. E. and Devereux, P. J. (2010). Recent developments in intergenerational mobility. *NBER Working Paper No. 15889*, pages 1–91.

Blanden, J. (2013). cross-country rankings in intergenerational mobility: A comparison of approaches from economics and sociology. *Journal of Economic Surveys*, 27(1):38–73.

Blanden, J., Gregg, P., and Macmillan, L. (2007). CE Accounting for Intergenerational Incom Persistence : Non- Cognitive Skills, Ability and Education. 117:CEE DP 73 ISSN 2045–6557.

Blanden, J., Haveman, R., Smeeding, T., and Wilson, K. (2014). Intergenerational mobility in the united states and great britain: A comparative study of parent-child pathways. *Review of Income and Wealth*, 60(3):425–449.

Boserup, S. H., Kopczuk, W., and Kreiner, C. T. (2014). Stability and persistence of intergenerational wealth formation : Evidence from Danish wealth records of three generations. *Unpublished Working Paper*.

Boserup, S. H., Kopczuk, W., and Kreiner, C. T. (2016). Born with a Silver Spoon? Danish Evidence on Wealth Inequality in Childhood. *NBER Working Paper Series*, (22549).

Bowles, S. (2004). *Microeconomics - Behavior, Institutions and Evolution*. Princeton University Press.

Bowles, S. (2012). *The New Economics of Inequality and Redistribution*. Cambridge University Press.

Bowles, S. and Gintis, H. (2002). Schooling in Captalist America Revisited. *Sociology of Education*, 75(1):1–18.

Bowles, S. and Gintis, H. (2011). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. Haymarket Books, Chicago, 2 edition.

Bowles, S., Gintis, H., and Groves, M. O. (2005). Introduction. In Bowles, S., Gintis, H., and Osborne Groves, M., editors, *Unequal Chances - Family Background and Economic*

Success, chapter 1, pages 1–22. Russell Sage Foundation & Prince ton University Press, New York and Princeton and Oxford.

Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., and Österbacka, E. (2007a). Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons. *The Economic Journal*, 117:72–92.

Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., and Österbacka, E. (2007b). Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons. *Economic Journal*, 117:72–92.

Chetty, R. (2016). the Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *NBER Working Paper No. 23001.*

Chetty, R. and Hendren, N. (2016). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *NBER Working Paper No. 23002*.

Chetty, R., Hendren, N., Jones, M., and Porter, S. (2018). Race and Economic Opportunity in the United States: An Intergenerational Perspective. *NBER Working Paper No. 24441.*

Chetty, R., Hendren, N., and Katz, L. F. (2016a). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Chetty, R., Hendren, N., and Katz, L. F. (2016b). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):90.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the Land of Opportunity? The Geopgraphy of Intergenerational Mobility in the United States. *Quarterly Journal of Economics*, 129(4):1553–1623.

Chiang, A. C. and Wainwright, K. (2005). *Fundamental Methods of Mathematical Economics*. McGraw-Hill, internatio edition.

Corak, M. (2004). *Generational Income Mobility in North America and Europe*. Cambridge University Press.

Corak, M. (2006a). Do Poor Children Become Poor Adults? Lessons from a Cross-Country Comparison of Generational Earnings Mobility. *Research on Economic Inequality*, 13(1993):143–188.

Corak, M. (2006b). Do poor children become poor adults? Lessons from a cross country comparison of generational income mobility. *IZA Discussion Paper*, (1993).

Corak, M. (2013). Inequality from Generation to Generation: The United States in Comparison. *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, (9929):107–126.

Corak, M. (2016). Inequality from Generation to Generation: The United States in Comparison. *IZA Discussion Paper*, (9929).

Corak, M., Lindquist, M. J., and Mazumder, B. (2014). A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics*, 30:185–200.

Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. *The American Economic Review*, 97(2):31–47.

Currie, J. (2001). Early Childhood Education Programs. *The Journal of Economic Perspectives*, 15(2):213–238.

Danmarks Statistik (2017a). Ækvivaleret Disponibel Indkomst. *Technical Document, July.*

Danmarks Statistik (2017b). Ulighedsmål: Gini-koefficient. *Technical Document, July.*

Del Boca, D., Flinn, C., and Wiswall, M. (2014). Houshold choices and child development. *The Review of Economic Studies*, 81(1):137–185.

Donnelly, L., Garfinkel, I., Brooks-Gunn, J., Wagner, B. G., James, S., and McLanahan, S. (2017). Geography of intergenerational mobility and child development. *Proceedings of the National Academy of Sciences*, 114(35):201700945.

DR (2016). *Tema: En Syg Forskel.* Danmarks Radio, Website, accessed on April 15, 2018, URL: https://www.dr.dk/nyheder/tema/en-syg-forskel.

Durant, W. (2006). *The Story of Philosophy - The Lives and Opinions of the Greater Philosophers*. Pocket Books, New York, N.Y., paperback edition.

Durlauf, S. N. and Mogstad, M. (2017). Intergenerational Mobility : Guest Editors ' Preface. *The Scandinavian Journal of Economics*, 119(1):3–4.

Eika, L., Mogstad, M., and Zafar, B. (2016). Educational Assortative Mating and Household Income Inequality. *Federal Reserve Bank of New York ,Staff Reports No. 682*, (682).

Eriksson, T., Bratsberg, B., and Raaum, O. (2005). Earnings persistence across generations: Transmission through health? *Department of Economics, University of Oslo, Memorandum No. 35*.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism*. Princeton University Press.

Fehr, E. and Schmidt, K. M. (2006). The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories. In Kolm, S.-C. and Ythier, J. M., editors, *Handbook of the Economics of Giving, Altruism and Reciprocity*, volume 1, chapter 8, pages 616–691. Elsevier.

Feldman, N. E. (2010). Mental Accounting Effects of Income Tax Shifting. *Review of Economics and Statistics*, 92(1):70–86.

Fleurbaey, M. and Peragine, V. (2013). Ex Ante Versus Ex Post Equality of Opportunity. *Economica*, 80(317):118–130.

Francesconi, M. and Heckman, J. J. (2016). Child Development and Parental Investment: Introduction. *Economic Journal*, 126(596):F1–F27.

Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton University Press.

Frisch, R. and Waugh, F. V. (1933). Partial Time Regressions as Compared with Individual Trends. *Econometrica*, 1(4):387–401.

Gelman, A. and Loken, E. (2014). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Psychological bulletin*, 140(5):1272–1280.

Gennetian, L. A., Sanbonmatsu, L., and Ludwig, J. (2011). An Overview of Moving to Opportunity A Random Assignment Housing Mobility Study in Five U . S . Cities. In Newburger, H. B., Birch, E. L., and Wachter, S. M., editors, *Neighborhood and Life Chances: How Place Matters in Modern America*, chapter 10, pages 163–178. University of Pennsylvania Press.

Grawe, N. D. (2006). Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics*, 13(5):551–570.

Greene, W. H. (2011). *Econometric Analysis*. Pearson, 7 edition.

Haider, S. and Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. *The American Economic Review*, 96(4):1308–1320.

Hall, P. A. and Soskice, D. (2001). An Introduction to Varieties of Capitalism. In Hall, P. A. and Soskice, D., editors, *Varieties of capitalism - the institutional foundations of comparative advantage*, chapter 1. Oxford University Press, Oxford.

Harding, D. (2003). Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy. *American Journal of Sociology*, 109(3):676–719.

Heckman, J. J. (2008a). Role of income and family influence on child outcomes. *Annals of the New York Academy of Sciences*, 1136:307–323.

Heckman, J. J. (2008b). Schools, skills, and synapses. *Economic Inquiry*, 46(3):289–324.

Heckman, J. J., Pinto, R., and Savelyev, P. (2013). American Economic Association Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *The American economic review*, 103(6):2052–2086.

Hertz, T. (2006). Understanding mobility in America. *Center for American Progress Discussion Paper, April 26, 2006*.

Hussain, M. A., Munk, M. D., and Bonke, J. (2009). Intergenerational Earnings Mobilities–How Sensitive Are They to Income Measures? *Journal of Income Distribution*, 18(3-4):79–92.

Jäntti, M., Bratsberg, B., Røed, K., Raaum, O., Österbacka, E., Björklund, A., and Eriksson, T. (2006). American exceptionalism in a new light : a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. *IZA Discussion Paper*, (1938):1–40.

Jencks, C. and Mayer, S. E. (1990). The Social Consequences of Growing Up in a Poor Neighborhood. *Inner-City Poverty in the United States*, pages 111–186.

Jones, R. (1976). *The Third Wave*. The Wave Home, Accessed on March 18, 2018 - URL: http://www.thewavehome.com/1976_The-Third-Wave_story.htm.

Kanbur, R. and Stiglitz, J. E. (2016). Dynastic inequality, mobility and equality of opportunity. *Journal of Economic Inequality*, 14(4):419–434.

King, G. and Nielsen, R. (2016). Why Propensity Score Should Not Be Used For Matching. *Unpublished Manuscript,*.

Landersø, R. and Heckman, J. J. (2016a). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S. *NBER Working Paper Series*, 22465.

Landersø, R. and Heckman, J. J. (2016b). Web Appendix for Scandinavian Fantasy : The Sources of Intergenerational Mobility in Denmark List of Figures. *NBER Working Paper Series - Online Appendix*.

Landersø, R. and Heckman, J. J. (2017). The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US. *Scandinavian Journal of Economics*, 119(1):178–230.

Leventhal, T. and Brooks-Gunn, J. (2000). The Neighborhoods They Live in: The Effects of Neighborhood Residence on Child and Adolescent Outcomes. *Psychological Bulletin*, 126(2):309–337.

Leventhal, T. and Brooks-Gunn, J. (2003). Moving to Opportunity: an Experimental Study of Neighborhood Effects on Mental Health. *American Journal of Public Health*, 93(9):1576–1582.

Ludwig, J., Liebman, J., Kling, J., Duncan, G., Katz, L., Kessler, R., and Sanbonmatsu, L. (2008). What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment? *American Journal of Sociology*, 114(1):144–188.

Marx, K. and Engels, F. (1978). Manifesto of the Communist Party. In Tucker, Robert, C., editor, *The Marx-Engels Reader*, chapter Part 3, 1, pages 469–500. W. W. Norton & Company, New York, N.Y., 2 edition.

Mayer, S. E. and Lopoo, L. M. (2008). Government spending and intergenerational mobility. *Journal of Public Economics*, 92(1-2):139–158.

Mazumder, B. (2005). Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data. *Review of Economics and Statistics*, 87(2):235–255.

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21(3):685–701.

Mazumder, B. (2011). Black-White Differences in Intergenerational Economic Mobility in the US. *Federal Reserve Bank of Chicago Working Paper No. 2011-10.*

Mazumder, B. (2016). Estimating the Intergenerational Elasticity and Rank Association in the United States: Overcoming the Limitations of Tax Data. In Capellari, L., Polachek, S. W., and Tatsiramos, K., editors, *Inequality: Causes and Consequences*, pages 83–129. Emerald Group Publishing Limited.

Mogstad, M. (2017). The Human Capital Approach to Intergenerational Mobility. *Journal of Political Economy*, 125(6):1862–1868.

Munk, M. D., Bonke, J., and Hussain, M. A. (2016). Intergenerational top income persistence: Denmark half the size of Sweden. *Economics Letters*, 140:31–33.

Nybom, M. and Stuhler, J. (2016). Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation. *Journal of Human Resources*, 51(1):239–268.

NYCHA (2018). *About Section 8.* New York City Housing Authority, Website, accesed on April 19, 2018 at URL: http://www1.nyc.gov/site/nycha/section-8/about-section-8.page.

O'Flaherty, B. (2015). *The Economics of Race in the United States*. Harvard University Press, Cambridge, Massachusetts & London, England.

Oreopoulos, P. (2003). The Long-Run Consequences of Living in a Poor Neighborhood. *The Quarterly Journal of Economics*, 118(4):1533–1575.

Oreopoulos, P., Page, M., and Stevens, A. H. (2008). The Intergenerational Effects of Worker Displacement. *Journal of Labor Economics*, 26(3):455–483.

Osborne Groves, M. (2005). Personality and the Intergenerational Transmission of Economic Status. In Bowles, S., Gintis, H., and Osborne Groves, M., editors, *Unequal Chances - Family Background and Economic Success*, chapter 7, pages 208–231. Russell Sage Foundation & Princeton University Press, New York and Princeton and Oxford.

Page, M. E. and Solon, G. (2003). Correlations between Brothers and Neighboring Boys in Their Adult Earnings: The Importance of Being Urban. *Journal of Labor Economics*, 21(4):831–855.

Pekkarinen, T., Uusitalo, R., and Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, 93(7-8):965–973.

Piketty, T. (2014). *Capital In The Twenty-First Century*. The Belknap Press of Harvard University Press, Cambridge Massachusetts.

PSC (2018). *Racial Residential Segregation Measurement Project*. Population Studies Center, University of Michigan, Population Studies Center, University of Michigan, Website, accessed on April 20, 2018, URL: http://enceladus.isr.umich.edu/race/calculate.html.

Putnam, R. (2015). *Our Kids - The American Dream in Crisis*. Simon & Schuster, New York, 1 edition.

Raaum, O., Salvanes, K. G., and Sørensen, E. (2006). The Neighbourhood Is Not What It Used To Be. *Economic Journal*, 116(508):200–222.

Rawls, J. (1999). *A Theory of Justice*. Belknap Press, Cambridge, MA, revised edition.

Regeringen (2007). *Lige muligheder - Styrkede personlige ressourcer og social sammenhængskraft*. Number September. The Ministry of the State, Accessed on March 2, 2018 - URL: http://www.stm.dk/multimedia/Regeringsudspil_om_Lige_muligheder.pdf.

Richey, J. and Rosburg, A. (2018). Decomposing economic mobility transition matrices. *Journal of Applied Econometrics*, 33:91–108.

Roemer, J. E. (1998). *Equality of Opportunity*. Harvard University Press, Cambridge, Massachusetts.

Roemer, J. E. (2000). Equality of Opportunity. In Arrow, K. J., Bowles, S., and Durlauf, S. N., editors, *Meritocracy and Economic Inequality*, pages 17–32. Princeton University Press.

Roemer, J. E. (2002). Equality of opportunity : A progress report. *Social Choice and Welfare*, 19(2):455–471.

Roemer, J. E. (2004). Equal Opportunity and Intergenerational Mobility: Going beyond Intergenerational Income Transition Matrices. In Corak, editor, *Generational Income Mobility in North America and Europe*, chapter 3. Cambridge University Press, Cambridge.

Rothbaum, J. (2016). Sorting and Geographic Variation in Intergenerational Mobility. *Unpublished Manuscipt*, (December):1–79.

Rothwell, J. T. and Massey, D. S. (2015a). Geographic Effects on Intergenerational Income Mobility. *Economic Geography*, 91(1):83–106.

Rothwell, J. T. and Massey, D. S. (2015b). Geographic Effects on Intergenerational Income Mobility. *Economic Geography*, 91(1):83–106.

Sacerdote, B. (2007). How large are the effects from changes in family environment? A Study of Korean American Adoptees. *The Quarterly Journal of Economics*, 122(1):119–157.

Sampson, R. (2008). Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure. *American Journal of Sociology*, 114(1):189–231.

Sampson, R. J. (2012). *Great American City: Chicago and the Enruding Neighborhood Effect.* Chicago University Press, Chicago.

Sandel, M. (2010). *Justice - What's The Right Thing To Do.* Penguin Books.

Schnitzlein, D. D. (2014). How important is the family? Evidence from sibling correlations in permanent earnings in the USA, Germany, and Denmark. *Journal of Population Economics*, 27(1):69–89.

Sen, A. (2006). Development as Freedom. *Indian Journal of Industrial Relation*, 42(2):157–169.

Sharkey, P. (2008). The Intergenerational Transmission of Context. *American Journal of Sociology*, 113(4):931–969.

Sharkey, P. and Torrats-Espinosa, G. (2017). The effect of violent crime on economic mobility. *Journal of Urban Economics*, 102:22–33.

Shefrin, H. M. and Thaler, R. H. (1988). The Behavioral Life Cycle Hypothesis. *Economic Inquiry*, 26(4):609–643.

Shorrocks, A. F. (1978). The Measurement of Mobility. *Econometrica*, 46(5):1013–1024.

Simon, H. A. (1992). Altruism and Economics. *Eastern Economic Journal*, 18(1):73–83.

Small, M. L. and Newman, K. (2001). Urban poverty after the truly disadvantaged: the rediscovery of the family, the neighborhood, and culture. *Annual Review of Sociology*, 27(2001):23–45.

Solon, G. (1989). Biases in the Estimation of Intergenerational Earnings Correlations. *The Review of Economics and Statistics*, 71(1):172–174.

Solon, G. (1992). Intergenerational Income Mobility. *American Economic Review*, (June):393–408.

Solon, G. (1999). Intergenerational Mobility in the Labor Market. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics, Volume 3, Part A*, chapter 29, pages 1761–1800. Elsevier Science B.V.

Solon, G. (2002). Cross-country Differences in intergenerational earnings mobility. *Journal of Economic*, 16(3):59–66.

Solon, G. (2004). A Model of Intergenerational Mobility Variation over Time and Place. In Cor, editor, *Generational Income Mobility in North America and Europe*, chapter 2. Cambridge University Press, Cambridge.

Solon, G. (2014). Theoretical models of inequality transmission across multiple generations. *Research in Social Stratification and Mobility*, 35:13–18.

Thakral, N. and Tô, L. T. (2017). *Daily Labor Supply and Adaptive Reference Points*. Unpublished Manuscript, available at http://scholar.harvard.edu/files/nthakral/files/ThakralTo_taxi.pdf.

Thaler, R. H. (2016). *Misbehaving - The Making of Behavioral Economics*. Penguin Books.

The Equality of Opportunity Project (2018). *Data and Replication Code*. Website, accessed on April 25, http://www.equality-of-opportunity.org/data/.

Tokey, J. W. (1980). We Need Both Explanatory and Confirmatory. *The American Statistician*, 34(1):23–25.

Torche, F. (2015). Analyses of Intergenerational Mobility: An Interdisciplinary Review. *The ANNALS of the American Academy of Political and Social Science*, 657(1):37–62.

UVM (2018). *Parallelsamfundsudspil skal sikre børn og unge lige muligheder*. Ministry of Education, Accessed on March 2, 2018 - URL: https://uvm.dk/aktuelt/nyheder/uvm/2018/mar/180301-parallelsamfundsudspil-skal-sikre-boern-og-unge-lige-muligheder.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, 1 edition.

Willis, R. J. (1986). Comment on Becker and Tomes. *Journal of Labor Economics*, 4(3(2)):40–47.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago University Press, Chicago.

Wilson, W. J. (1991). Another Look at The Truly Disadvantaged. *Political Science Quarterly*, 106(4):639–656.

Wodtke, G. T., Elwert, F., and Harding, D. J. (2016). Neighborhood Effect Heterogeneity by Family Income and Developmental Period. *American Journal of Sociology*, 121(4):1168–1222.

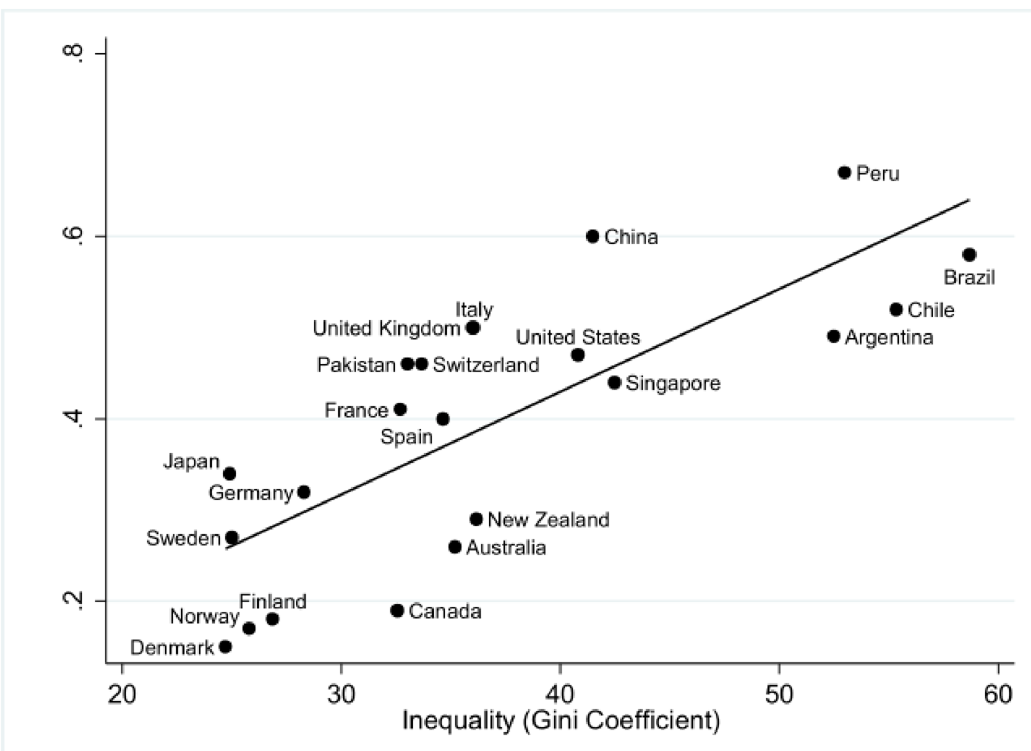
Appendices

Appendix A

Existing Research

A.1 The Great Gatsby Curve

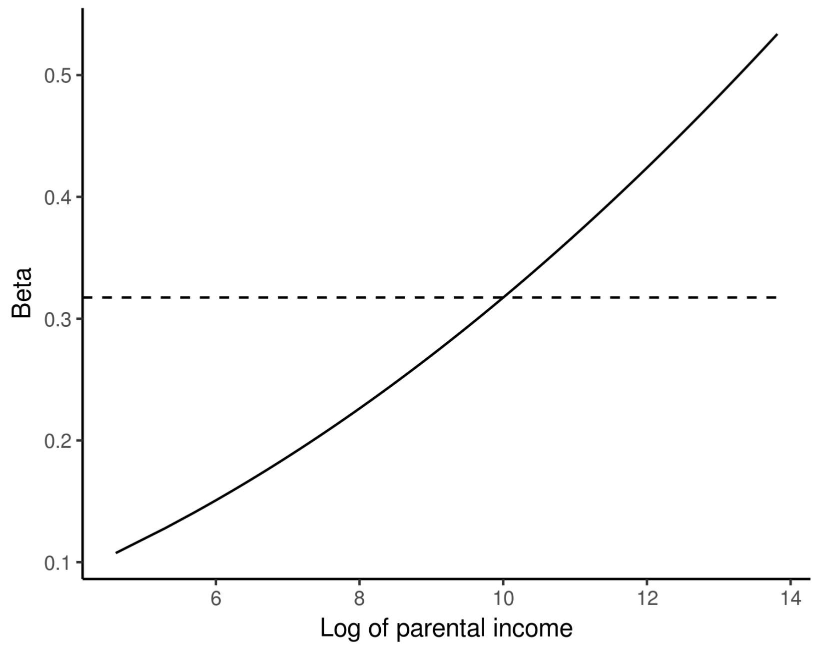
Figure A.1: "The Great Gatsby Curve" shows the correlation between measures of intergenerational father son earnings elasticities and inequality as measured by Gini Coefficients. Source (Corak, 2016, figure 2)



Note: Estimates of intergenerational mobility are collected by Corak (2016), and gini coefficients are collected from the World Bank (no identification of year of measurement is provided). The line shows a linear regressions.

A.2 Non-linearities in IGEs

Figure A.2: β coefficients for income varying and constant heritability of traits in the Solon (2004) model.



Note: The figure shows β values by parental income level. The full line indicates a heritability parameter that varies positively with parental income, while the dashed line indicates findings for a constant heritability across the income distribution. Parameter values are $\theta = 0.2$, $\gamma = 0.2$, $p = 0.12$. The constant λ is set to 0.3, while the income-varying parental income is $\lambda_{inc} = 0.01 \log y_{\text{i,t-1}} + 0.002 \log y_{\text{i,t-1}}$.

A.3 Derivation for Solon (2004)

For the model explained in section CHECK, the following is the derivation of the expression for the relation between parental and child income in equation 3.11.

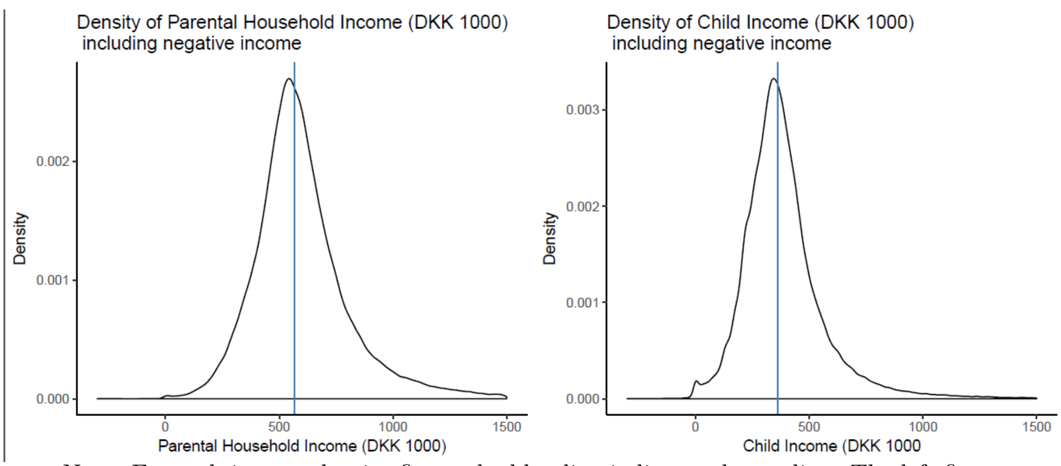
$$
\begin{aligned}
\log y_{\text{i,t-1}} &= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}} - \left[\frac{1-\alpha}{1-\alpha(1w-\theta p)}\right]G_{\text{i,t-1}} + G_{\text{i,t-1}}\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}} - \left[\frac{1-\alpha}{1-\alpha(1-\theta p)}\right]G_{\text{i,t-1}} + \left[\frac{1-\alpha(1-\theta p)}{1-\alpha(1-\theta p)}\right]G_{\text{i,t-1}}\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}} + \left[\frac{1-\alpha(1-\theta p)-(1-\alpha)}{1-\alpha(1-\theta p)}\right]G_{\text{i,t-1}}\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}} + \left[\frac{\alpha\theta p}{1-\alpha(1-\theta p)}\right]G_{\text{i,t-1}}\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right][(1-\tau)y_{\text{i,t-1}} + G_{\text{i,t-1}}]\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)y_{\text{i,t-1}}\left[1+\frac{G_{\text{i,t-1}}}{(1-\tau)y_{\text{i,t-1}}}\right]\right) + e_{\text{i,t-1}}\right] \\
&= \mu + p\left[\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)\right) + \log\left(y_{\text{i,t-1}}\right) + \log\left(1+\frac{G_{\text{i,t-1}}}{(1-\tau)y_{\text{i,t-1}}}\right) + e_{\text{i,t-1}}\right] \\
&= \mu + \underbrace{p\theta \log\left(\left[\frac{\alpha\theta p}{-\alpha(1-\theta p)}\right](1-\tau)\right)}_{A} + p\log\left(y_{\text{i,t-1}}\right) + \underbrace{p\log\left(1+\frac{G_{\text{i,t-1}}}{(1-\tau)y_{\text{i,t-1}}}\right)}_{B} + pe_{\text{i,t-1}}
\end{aligned}
\tag{A.1}
$$

Appendix B

Income

B.1 Income Densities

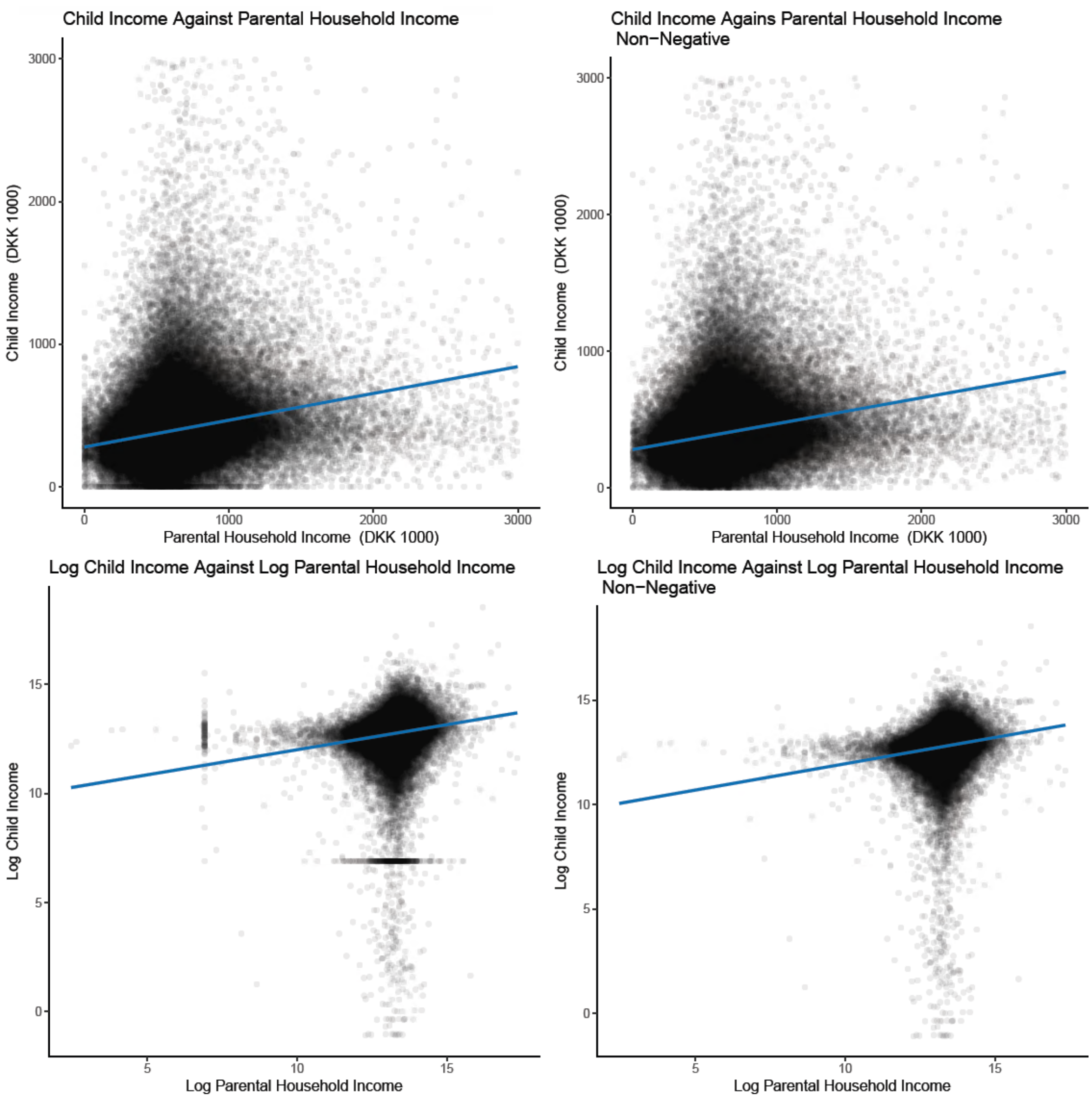
Figure B.1: Density plots of parental household and child total income including negative income.



Note: For each income density figure the blue line indicates the median. The left figure shows the density of total parental income in 2015 (1000 DKK) including negative income observations. Total income includes labor earnings, income from own business, capital income, and public transfers. Parental income is measured during the period when the children are at the age of 7-15. The right figure shows similar total income for children measured between the ages 35-37, 36-38, and 37-39 for children born in 1975, 1974, and 1973 respectively.

B.2 Income Scatter plots

Figure B.2: Scatterplot of parental household and child total income and log of total income.



Note: Blue lines indicate simple linear regressions lines. The figures show (top-left) total child income against total parental household income with recoded non-positive income observations. (Lower-left) total child income against total parental household income removing non-positive income observations. (Upper-right) Log total child income against log total parental income with recoded non-positive income. (Lower-right) Log total child income against log total parental household income with removed non-positive observations. Slopes of the regressions lines in the upper-right and lower-right plots show the estimated IGE at the national level at .097 and .291, respectively.

B.3 Income Percentiles for the Total Sample

Table B.1: Total income and sample sizes for children's, parental households', mothers', and fathers' total income

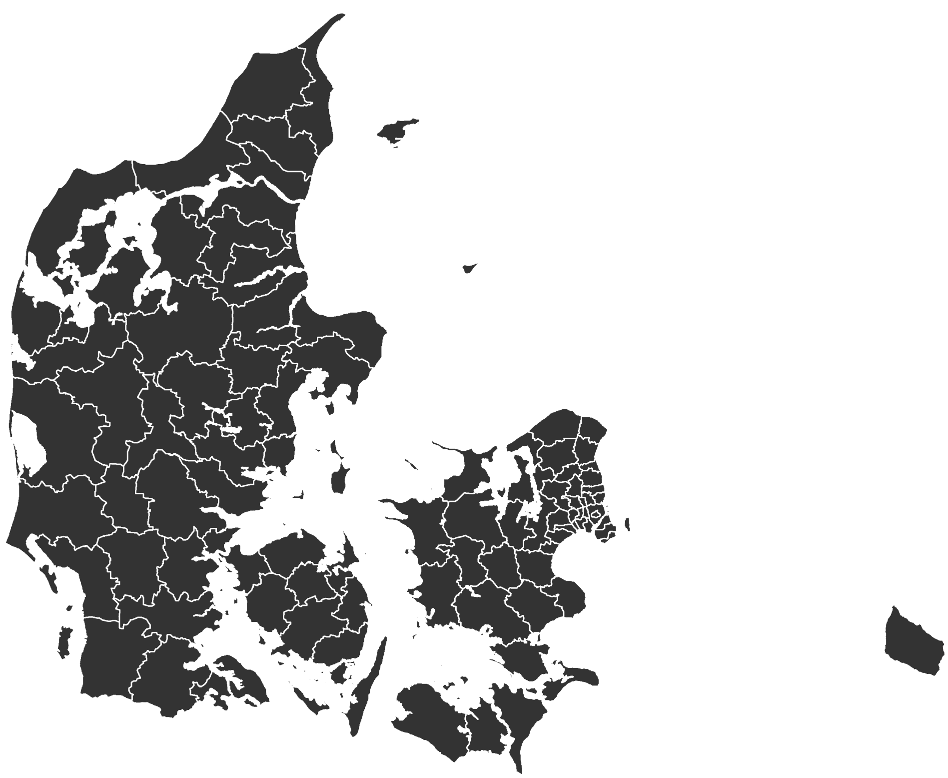
	5th	10th	20th	30th	40th	50th	60th	70th	80th	90th	95th
Child	160.71	207.85	260.43	300.40	332.68	362.82	395.07	433.77	488.18	593.87	721.76
Parent Household	292.16	359.57	440.78	491.59	532.01	570.20	611.24	661.36	732.50	867.43	1,038.22
Mother	29.49	81.20	136.95	168.25	193.38	213.30	232.20	254.60	285.06	331.69	371.78
Father	136.72	207.28	269.25	299.58	325.36	351.95	382.74	423.65	486.72	608.49	763.80

Note: The table contains information on the total income measure of all individuals in the main sample used in chapter 3 for percentiles 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95. Total income is the sum of labor income, personal business income, income from financial wealth, and all transfers.

Appendix C

Danish Municipalities

Figure C.1: Danish municipalities following the 2007 administrative reform.

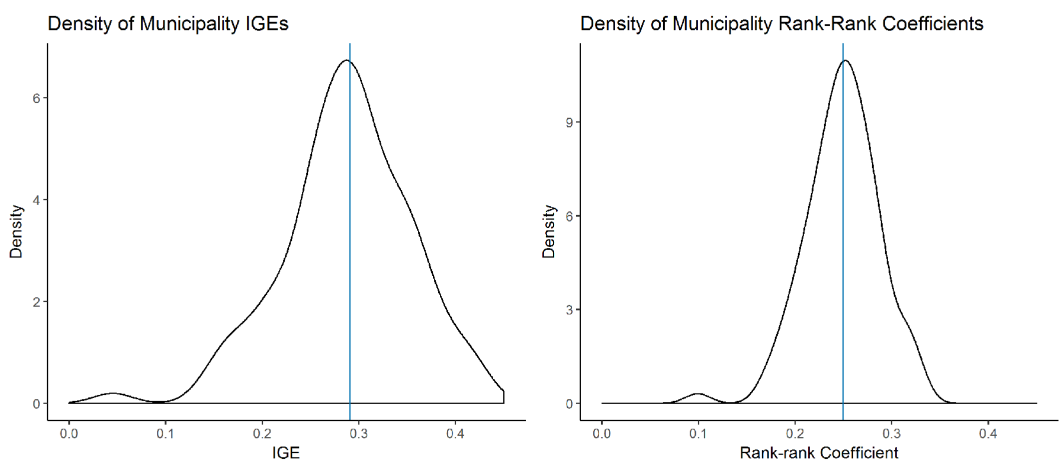


Appendix D

Municipality Results

D.1 Densities of IGE and Rank-Rank Coefficients

Figure D.1: Densities of Municipality IGEs and Rank-rank Coefficients.

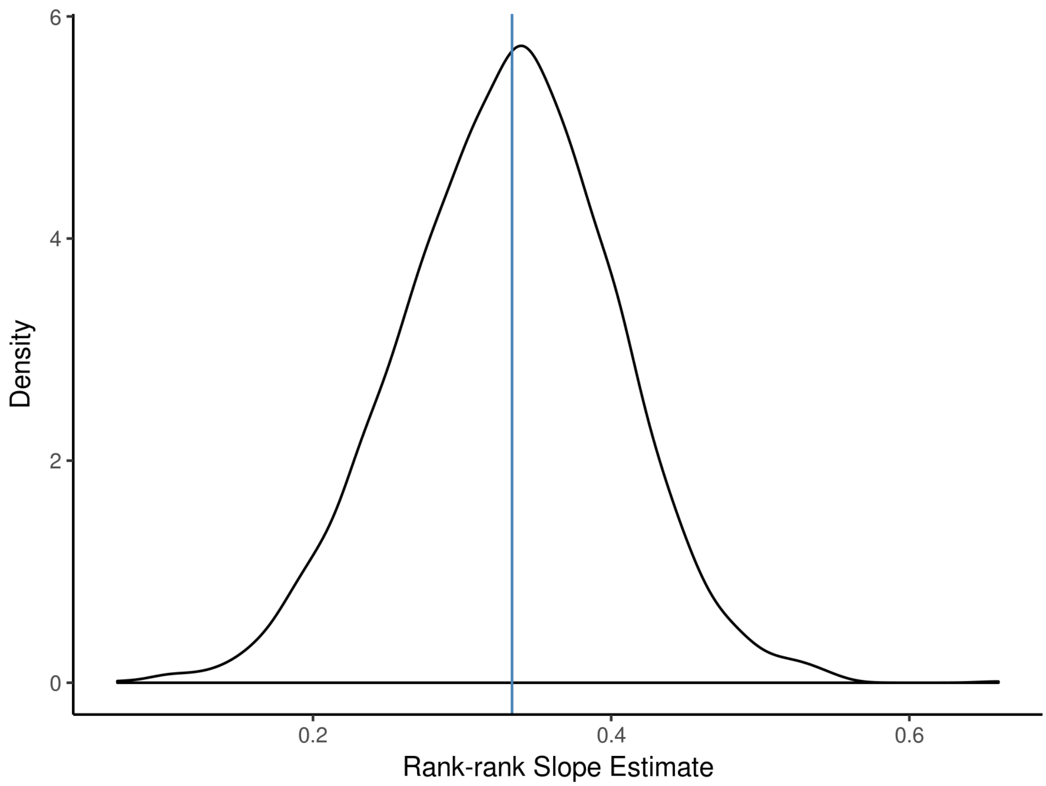


Note: The left-hand figure shows the density of IGEs estimated for the 98 Danish municipalities. The IGE is calculated as the regression coefficient between log child total income and log parental household total income, dropping non-positive observations. The Right-hand side figure shows the density of the rank-rank coefficients from a regression of the child income rank (within cohort) against parental income rank (within child cohort).

D.2 County Mobility from Chetty et al. (2014)

The findings in this section are based on the work by Chetty et al. (2014). The summary statistics and mobility estimates at the county level in the US is available online of their findings are available online at The Equality of Opportunity Project (2018). In figure D.2 I show a density plot of the US rank-rank coefficients. Comparing the plot to the Danish RRC estimates in chapter 3, it is clear that there is both a larger variation in the estimated coefficients, and that the average relation between parent and child rank is stronger in the US (for counties) than in Denmark (for municipalities).

Figure D.2: Densities of Conuty Rank-Rank Coefficients from Chetty et al. (2014).



Note: The figure shows a density plot of US county Rank-rank regression coefficients based on the work by Chetty et al. (2014). The blue line indicates the median rank-rank coefficient.

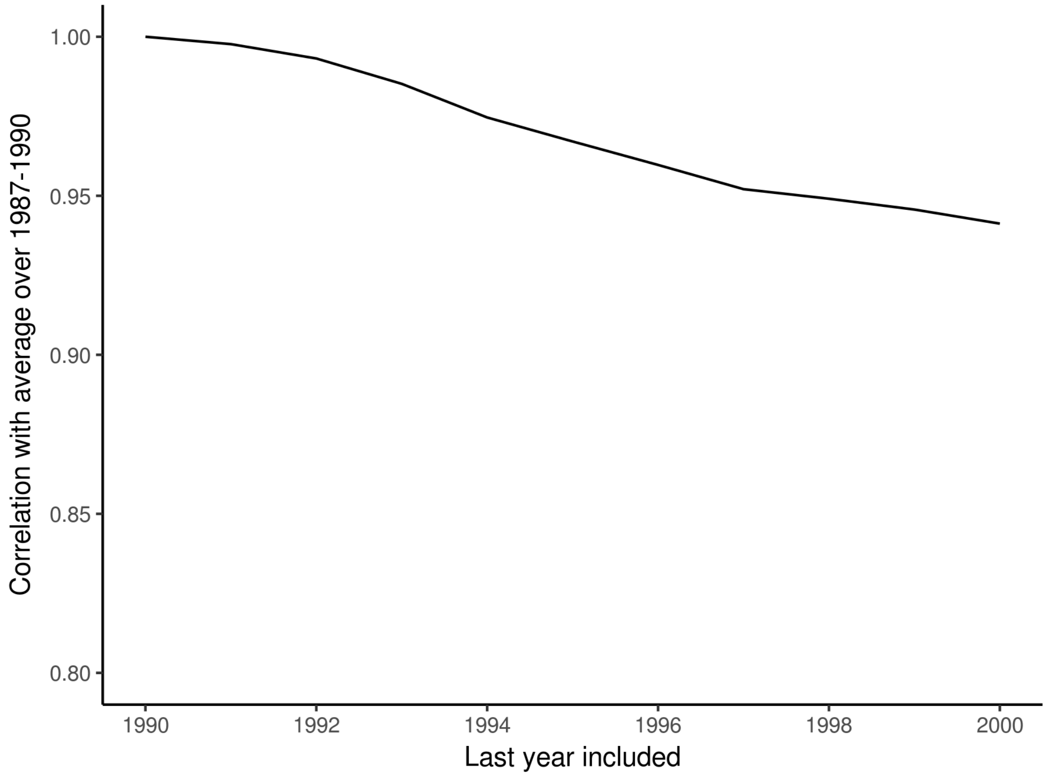
Appendix E

Neighborhood Effects

E.1 Gini-coefficient

The gini coefficients used in chapter 4 are collected from Statistics Denmark. The estimates are only available from 1987 and onwards. I utilize an average over the period 1987-1993 in the text to smooth out idiosyncratic variations in within municipality inequality. Figure E.1 shows the correlation in estimates of municipality income inequality by gini coefficients when extending the period of measurement.

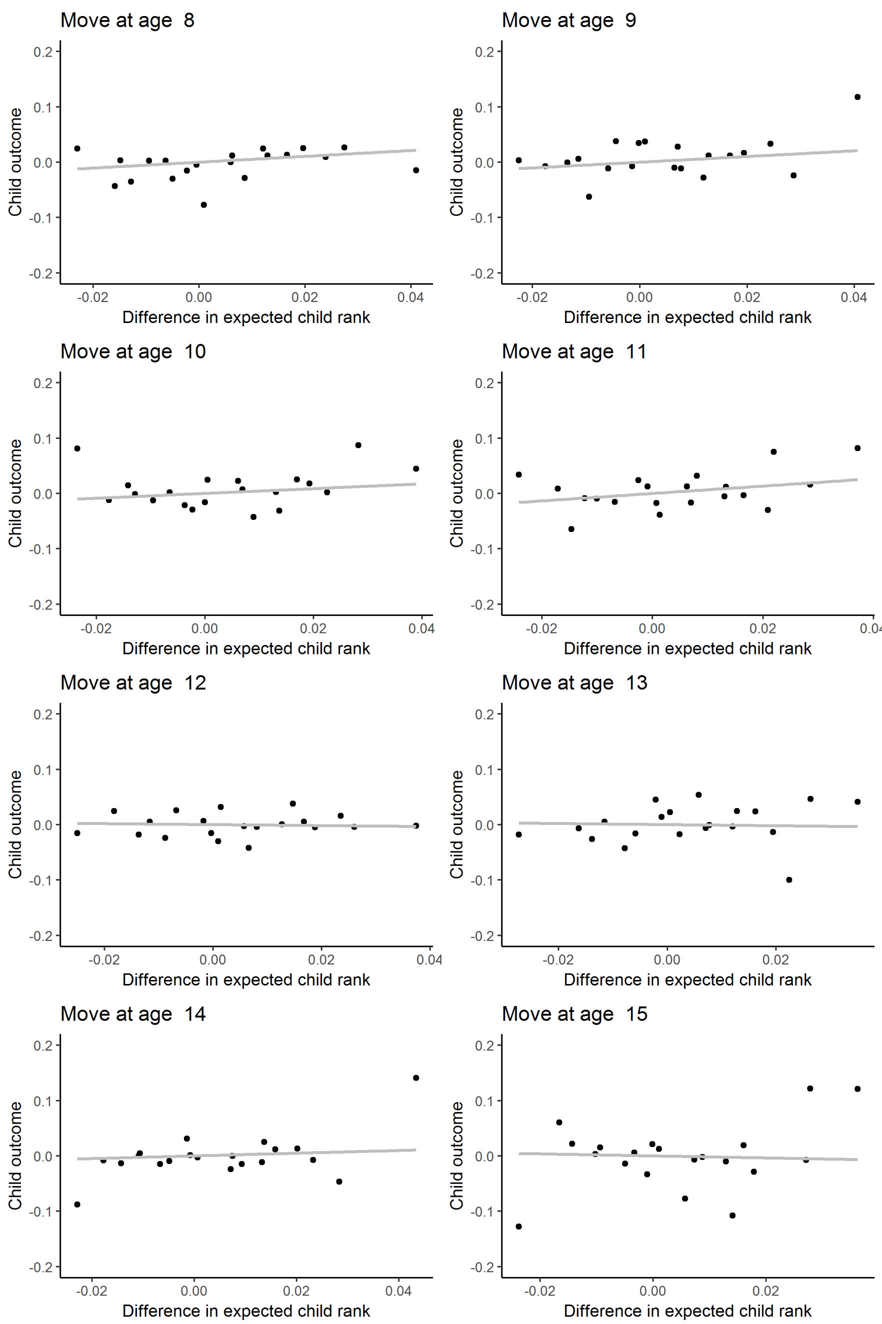
Figure E.1: Correlations in municipality gini coefficients when extending averaging period



Note: Correlation in among municipality level gini coefficients when adding years to the period of averaging. The baseline is 1987-1990. For each year the period is extended with, the graph shows the correlation with the baseline municipality averages.

E.2 Relations between the expected child rank and difference in outcome municipality

Figure E.2: Binned scatterplots for child income-rank given expected origin-outcome difference, controlling for origin municipality and family fixed effects, age 8-15.



Note: The plot shows binned scatterplot values. For each ventile of difference in origin-destination expected child rank, the average outcome of children in the age-group has been averaged. The values are calculated for each group of movers from age 8 to 15. The grey line shows indicates the 0 intercept.