School Choice, School Performance and School Segregation

Institutions and Design

Dany Kessel

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Monday 3 September 2018 at 14.00 in De Geersalen, Geovetenskapens hus, Svante Arrhenius väg 14.

Abstract

This thesis consists of four self-contained chapters.

The first chapter, **Are Parents Uninformed? – The Impact of School Performance Information on School Choice Behavior and Student Assignment**, is co-authored by Elsisabet Olme. We investigate the effects of school performance information on school choice behavior and student assignment. A randomly selected group of students, about to choose middle school, were provided with information about the performance of the available schools. Households that received the information became more prone to choose a top-performing school. This effect is driven by native and high-skilled households. We simulate how this change in choice behavior translates into changes in school assignment. We find that enrollment in the top-performing schools increases but the effect is muted by limited capacity. We also find that the treatment increases the gap in school performance between advantaged and disadvantaged households, decreases segregation in terms of migration background and increases segregation in terms of parental skill-level.

The second chapter, **School Choice Priority Structures and School Segregation**, is also co-authored by Elsisabet Olme. We evaluate how school segregation is affected by altering the priority structures in a school choice program. We evaluate three priority structures, one proximity-based, one lottery-based and one based on soft quotas. Using actual choice data and simulations we find that that priority structures do affect school segregation. When reserving seats for different groups, schools are less segregated compared to when using systems where priorities are based on proximity or a lottery. We find that the average costs in terms of welfare are limited but that the different priority structures benefit different subgroups.

In the third chapter, **Debiasing the Gender Differences in Willingness to Compete – The Effects of General Information on the Gender Gap and Efficiency**, I explore if informing people about the gender differences in the willingness to compete and the accompanying inefficiencies can reduce said differences and inefficiencies. In an experiment where the participants got to choose whether to compete or not, a random sample of participants were informed about the gender differences in willingness to compete and the related inefficiencies. Among those not informed, men were much more likely to compete than women. There were also significant inefficiencies from low-performing men choosing to compete and high-performing women choosing not to. The treatment reversed the gender gap and significantly reduced inefficiency.

The fourth chapter, **The Housing Wealth Effect: Quasi-Experimental Evidence** is co-authored by Roine Vestman and Björn Tyrefors Hinnerich. We exploit a quasi-experiment that occurred in Stockholm in 2007 when the contract of Stockholm's city airport was unexpectedly renewed. We estimate an immediate shock of approximately 16 percent to house prices close to the airport. This source of price variation is ideal to identify housing wealth effects since it is local and unrelated to variation in macroeconomic conditions. Using a household data set with granular geographic information on primary residence, we find an MPC on cars of less than 0.2 cents per dollar.

Keywords: *School choice, school segregation, school performance, priority structures, information, gender, competition, experiment, house prices, housing wealth, consumption, house price elasticity, marginal propensity to consume.*

Stockholm 2018
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-156309

ISBN 978-91-7797-334-8
ISBN 978-91-7797-335-5

Department of Economics

Stockholm University, 106 91 Stockholm

SCHOOL CHOICE, SCHOOL PERFORMANCE AND SCHOOL SEGREGATION: INSTITUTIONS AND DESIGN

Dany Kessel

Stockholm University

School Choice, School Performance and School Segregation: Institutions and Design

Dany Kessel

ISBN print 978-91-7797-334-8
ISBN PDF 978-91-7797-335-5
ISSN 1404-3491

Front page: Erik Rein-Hedin

Printed in Sweden by Universitetsservice US-AB, Stockholm 2018
Distributor: Department of Economics, Stockholm University

To whom it may concern

Acknowledgements

Jonas Vlachos, my main supervisor, have been an integral part of my thesis work. Without Jonas curiosity, general knowledge and ability to tell me when I was going off on tangents, my thesis would likely have been much more of a mess than it is today. I also want to thank my co-supervisor and co-author Roine Vestman who, more than anyone, has helped me understand the strange creature that is academia and my co-author, Björn Tyrefors Hinnerich who has been incredibly supporting and generous all though my Ph.D.

Elisabet Olme has, for the last five years has been my main collaborator both inside and outside of academia. I have a hard time imagining that I would have been able to complete this thesis without Elizabet's clear-sightedness, work ethic and ability to take on any task.

There are a large number of scholars at different institutions that have taken time out of their busy schedules to answer questions and provide feedback. These include, but are not limited to, Matz Dahlberg, Emma von Essen, Anna Dreber, Astri Muren, Karin Hederos, Anna Sandberg, David Strömberg, Per Petterson-Lidbom, Tommy Andersson, Anna Tompsett, Douglas Almond, William Greene, Peter Bergman, Chris Nielsen, Miguel Urqinola, Jonah Rockoff and Hendrik Wolff.

I would also like to thank a number of fellow Ph.D.-students that

have contributed to making this the most stimulating and fun time of my life: Jonas Cederlöf, Niklas Blomqvist, Erik Lindgren, Matilda Kilström, Sirus Dehdari, Julia Boguslaw, Erik öberg, Hannes Malmberg, Karl Harmenberg, Siri Isaksson, Adam Altmejd, Agneta Berge, Axel Cronert and Jacob Molinder. A special thanks to Elin Molin for always making me laugh and for always having my back!

Futher, I need to express my gratitude to a number of people that, in different ways, have helped me with various practical matters. Birgitta Odén and Klas Lind provided the data and the institutional knowledge used in my essays on school choice. Niklas Troedosn, Matilda Gullbransen Mattson, Arne Floryd and Johan Lundström allowed me access to their classrooms to conduct my experiments. Henrik Wallin and Mattias Hallberg helped me conduct said experiments. Lars Norén and Sasha Norén have proofread parts of this thesis. Erik Rein-Hedin and Emma Fastesson Lindgren have provided illustrations. I would also like to thank the administrative staff at the Department of Economics at Stockholm University that have been helpful in so many ways.

Last but definitely not least, I would like to thank my loving family that have been incredibly supportive and always provided me with a place of refuge when needed, my friends who I have spent far too little time with and most importantly, Leigh who has put up with, among other things, me working far too many late nights and spending a year on the other side of an ocean. I love you!

So long and thanks for all the fish!

Dany Kessel
Stockholm, Sweden
May 2018

Contents

Introduction

This thesis consists of four self-contained chapters. The first two chapters, after which this thesis is named, are both on the subject of school choice. This subject has been my main research interest over the last three years and is the foundation of my future research agenda. The overarching motivation for both these studies is the rapid and global expansion of school choice programs that we have seen over the last few decades. Scholars in various fields have, over the last decade or so, made great progress in trying to understand both the consequences of this development and how school choice programs, and the institutions surrounding them, optimally should be designed. There are however still a large number of questions where we do not have sufficient answers. Due to data limitations, this is especially true when it comes to questions where there is no clear theoretical answer. It is also clear that practice lags behind theory and I think that most scholars in this field would agree with me in the claim that the design of all school choice programs currently in use, has flaws. These two chapters can be thought of as my first contribution to the ongoing effort to try to correct this. The last two chapters are not related to school choice but rather projects I started before I chose to specialize in Economics of Education in general and school choice in particular.

The first chapter, **Are Parents Uninformed? - The Impact**

of School Performance Information on School Choice Behaviour and Student Assignment, is co-authored by Elisabet Olme. We investigate the effects of school performance information on school choice behavior and student assignment. The chapter is based on a large-scale field experiment that we conducted in a Swedish municipality in 2016. We randomized students about to choose middle school into a treatment group and a control group. Households in the treatment group received information about the available schools' absolute performance as well as performance adjusted for the schools' student composition. We do not find any significant effect on either the total share of students or the composition of students opting out of their default schools. We do however find that the share of households choosing a top-performing school increases by 5.2 percentage points (corresponding to 12%) due to the treatment. This effect is mainly driven by native and high-skilled households opting out of second-tier schools. Using the municipality's actual allocation mechanism, we simulate how this change in choice behavior translates into changes in school assignment. Enrollment in the top-performing schools increases but the effect is somewhat muted by the schools' capacity constraints. This increase is, again, driven by high-skilled and native households. These households do, to an extent, displace students with a foreign background in the top-performing schools. Combined, this leads to an increase in the gap in school performance between advantaged and disadvantaged households. This initial effect on assignment also propagates through the system. The students moving into the top-performing schools leave after them free seats at some of the oversubscribed middle-performing schools. These seats are, in large, filled by students with a foreign background. This, combined with the fact that the average share of students with a foreign background in the top-performing schools is quite high, leads to a 6-point decrease in school segregation (corre-

sponding to 14%) as measured by the Duncan Dissimilarity Index. Segregation in terms of parental skill-level does however seem to increase somewhat as students from high-skill households congregate in the top-performing schools.

The second chapter, **School Choice Priority Structures and School Segregation**, is also co-authored by Elisabet Olme. We evaluate how school segregation and student welfare is affected by altering the priority structures in a school choice program.[1] We use actual choice data from the Swedish municipality of Botkyrka to estimate parental preferences for school attributes. These preferences are, in turn, used in simulations of counterfactual allocations of students under different types of priority structures. We evaluate three different types of priority structures, one that is based on proximity to the school, one based on a lottery and one based on so called soft quotas or reserved seats. The results suggest that the priority structures do affect the level of school segregation. When reserving seats for different groups, schools become less segregated both in terms of parental education and foreign background compared to lottery- and proximity-based priority structures. We find that the average costs in terms of welfare are limited. Even though the average effects on welfare are small we do see that the different priority structures benefit different subgroups. The priority structure based on proximity, benefits highly-educated and native households, while the priority structure based on reserved seats, benefits highly-educated and foreign households and low educated and native households.

In the third chapter, **Debiasing the Gender Differences in Willingness to Compete - The Effects of General Information on the Gender Gap and Efficiency**, I explore if informing

[1]Together with the assignment mechanism, the priority structure determines which applicants are admitted to a school when the number of applicants exceed the number of available seats.

people about the well-documented gender differences in the willingness to compete and the accompanying inefficiencies can reduce said differences and inefficiencies. An experiment where the participants got to choose whether to compete or not was conducted in a Swedish high school. A random sample of participants were informed about the gender differences in willingness to compete and the related inefficiencies. Among those not informed, men were much more likely to compete than women. There were also significant inefficiencies from low- performing men choosing to compete and high-performing women choosing not to. The information treatment reversed the gender gap and significantly reduced inefficiency. The effect on inefficiency differs between groups with low-performing men and high-performing women being the winners and low-performing women the losers.

The fourth and final chapter, **The Housing Wealth Effect: Quasi-Experimental Evidence**, is co-authored by Roine Vestman and Björn Tyrefors Hinnerich. We exploit a quasi-experiment that occurred in Stockholm in 2007 when the contract of Stockholm's city airport (Bromma) was unexpectedly renewed after political negotiation behind closed doors. We estimate an immediate negative shock of approximately 16 percent to house prices close to the airport. This source of price variation is ideal to identify housing wealth effects since it is local and unrelated to variation in macroeconomic conditions. Using a household data set with granular geographic information on primary residence, we find an elasticity among purchasers of new cars of around 0.4 and an MPC of 2-3 cents per dollar for homeowners that buy a new car. Accounting for the fact that homeowners do not buy a car, the aggregate MPC on cars is less than 0.2 cents. This effect is considerably smaller compared to most previous esimates. We discuss a number of reasons for why this might be the case, including the fact that what we are estimating is likely closer to

a true partial equilibrium effect. The effect is concentrated to homeowners with a loan-to-value ratio between 0.6 and 0.8, consistent with a collateral effect.

Chapter 1

Are Parents Uninformed?

- The Impact of School Performance Information on School Choice Behaviour and Student Assignment*

*This chapter was co-authored by Elisabet Olme. We thank Matz Dahlberg, Jonas Vlachos and seminar participants at Stockholm University and Stockholm School of Economics for valuable comments and feedback. We also gratefully acknowledge financial support from *Ge alla elever samma chans* (*Ge alla elever samma chans* is a project that aims to reduce the inequality in the Swedich educational system. The project is run and funded by The Swedish Trade Union Confederation (LO) and the two major Swedish teachers unions (Lärarnas Riksförbund and Lärarförbundet)).

1.1 Introduction

Over the last three decades there has been a rapid and large expansion of school choice programs all over the world (Musset, 2012). Due to this, an increasing number of households are now faced with the decision of what school to send their children to. This expansion has been motivated by two main arguments. The first argument is that school choice will increase school quality. The mechanism in mind is that when parents get the opportunity to choose schools they will choose higher-performing schools. Schools will therefore have to improve their performance in order to attract students. The second argument is that school choice will allow students living in impoverished neighborhoods with failing schools, to opt out of these schools. Student sorting will become more based on ability and ambition and less on factors such as socioeconomic status or migration background. The empirical evidence on the effects of introducing or expanding school choice opportunities on school performance is mixed. Some studies find positive effects, others find no effects and a few even find negative effects.[2] Studies that have looked at the effects on school segregation all show that it increases when school choice is introduced or expanded.[3]

One reason why we have not seen the gains envisioned by the proponents of school choice could be a lack of easily accessible school performance information. For the market mechanism to work, households need to be able to observe school performance. As there is usually no price mechanism to guide households in the school market,

[2]See, for example, Hoxby (2000), Lavy (2010), Böhlmark and Lindahl (2015) and Lavy (2015) for studies finding positive effects, Cullen et al. (2006) and Hsieh and Urquiola (2006) for studies finding no effects Abdulkadiroglu et al. (2018) for a study finding negative effect.

[3]See Hsieh and Urquiola (2006), Bifulco and Ladd (2007), Söderström and Uusitalo (2010) and Böhlmark et al. (2016).

this can be a challenge. Hence, it is possible that easily accessible school performance information could strengthen the above described mechanism and increase the competitive pressure on schools. Further, the ability to observe school performance might not be equally distributed among households. There is evidence suggesting that households choose different types of schools depending on their socioeconomic status and migration background. More disadvantaged households seem to choose lower performing schools compared to more advantaged households.[4] It is unclear to what extent this is due to differences in underlying preferences between different types of households or to what extent it is caused by other differences such as, for example, differences in information available or in access to high-performing schools.[5] If more advantaged households face a lower cost of obtaining relevant information than more disadvantaged households then making school performance information more accessible could reduce the segregating effects of school choice.

To shed further light on these questions we conducted a large-scale randomized controlled trial in the Swedish municipality of Linköping. Households with children about to start middle school (7^{th} grade) in the autumn of 2016 were randomly assigned a treatment and a control group. Prior to the school choice period, treated households received a letter containing information on the absolute and adjusted (for student composition) performance of the available schools. Performance was measured by the test scores on standardized tests taken in 9^{th} grade.[6] Using this variation in access to school performance

[4]See Abdulkadiroglu et al. (2017), Abdulkadiroglu et al. (2018) and Hastings et al. (2009).

[5] Hastings and Weinstein (2008) show that lack of information seems to explain part of the choice behavior among disadvantaged households and Burgess et al. (2015) show that a significant proportion of the difference in school choice behavior between advantaged and disadvantaged households is due to differences in the quality of schools they have available to them.

[6]Prior to conducting the experiment we got approval from the Regional Ethical

information, we study the effects on school choice behavior. Further, we combine this random variation with the assignment mechanism used in Linköping to study how changes in choice behavior translates into changes in assignment, had all students been given the same information.

We are not the first to study this subject. A number of studies have tried to exploit school rankings from newspapers, school accountability schemes or government inspections, in order to explore the effects of school performance information. One approach to identifying the effects has been to exploit variation in the timing of when such information is made public. Hart and Figalo (2015) study the introduction of a school accountability scheme in Florida and find a 6 percent increase in enrollment in schools receiving an A compared to schools receiving a B or C. Hussain (2013) exploits the timing of school inspections in the UK and documents a 3 percent increase in enrollment the year after a school is rated as *outstanding* and a 4 percent decline in enrollment the year after a *fail* rating. Koning and van der Wiel (2013) analyze whether information on high school quality published by a national newspaper (Trouw) in the Netherlands affects enrollment. They find that enrollment in top-ranked academic tracks increases by almost 20 percent.[7] Similar studies have also been performed using housing prices as the outcome.[8]

Vetting Board in Stockholm (DNR 2015/1252-32).

[7]This effect seems large and as Ruijs and Oosterbeek (2017) point out this might be due to the fact that Koning and van der Wiel (2013) use enrollment in the third year of secondary school as a proxy for applications in the first year. In doing so they might not only pick up effects on initial enrollment but also effects on the proportion of students held back and/or transferring to other tracks. Using actual choice data from Amsterdam, Ruijs and Oosterbeek (2017) estimate a conditional logit model and find no evidence that these published school quality indicators predict choice behavior.

[8]Figalo and Lucas (2004) analyze the introduction of the school accountability scheme in Florida and document that house prices increase in schooling zones where the local school received a good score. Fiva and Kirkeboen (2011) study

Another approach used to identify the effects of these school rankings is to exploit different cutoffs that arise due to the design of the system. Mizala and Urquiola (2013) exploit the features of the SNED program in Chile which identifies high-performing schools among a set of similar schools in a sharp regression discontinuity design and find no effects on enrollment from being identified as an effective school. Lépine (2015) conducts a similar study in Brazil where results on standardized tests (the Enem) are published only for schools with a minimum number of test takers and finds no effects on enrollment of having test score results published for either high or low-performing schools.

The fact that the two approaches seem to yield different results raises the question if one, or both of them, fail to reproduce the results that a randomized controlled trial would produce. Given the limited number of studies this does not need to be the case. It could also be that, by chance, the studies exploiting cutoffs for identification were conducted in settings where there were no effects, while the studies that rely on the timing of the release of information were conducted in settings where there were effects. In fact, there are reasons to believe that the settings matter in these studies as they all rely on enrollment as their main outcome variable. Ruijs and Oosterbeek (2017) point out that effects on enrollment are not easily translated into effects on choice behavior. For example, finding no effects on enrollment does not necessarily imply that there were no effects on school choice behavior. Households may react to the information but

the release of previously unpublished information on school quality in Norway and find that housing prices are affected by this information. Imberman and Lovenheim (2016) exploit the release of value-added measures in Los Angles and find no effect on enrollment. The effects in both Figalo and Lucas (2004) and Fiva and Kirkeboen (2011) are short-lived reverting back to the pre-information prices in just a few months. This is not easily explained and could indicate that the results are, at least in part, driven by some sort of speculative behavior from house buyers.

institutional factors (such as capacity constraints or the assignment mechanism) may limit the extent to which this is translated into effects on assignment. Further, an effect on enrollment only tells us that there is an effect on choice behavior, not how large it is. Small changes in choice behavior can propagate through the system and create big changes in student assignment and big changes in choice behavior can be muted by institutional factors.

This means that in order to be able to predict the effects of school performance information interventions we need actual estimates of the effects on choice behavior. Due to data limitations there have not been many studies on the effects of school performance information when it comes to actual choice behavior. There are two main exceptions. The first is Hastings and Weinstein (2008) who use two experiments (one field experiment and one natural experiment which arose because of the design of the No Child Left Behind Act) that provided direct information about school test scores in lower-income areas in the Charlotte-Mecklenburg Public School District. Using school choice data, they document an 5 percentage point increase in the probability of opting out from one's neighborhood school and instead choosing a higher-performing school. The second is Corcoran et al. (2018) who randomly select a number of high-poverty New York City middle schools and provide the students in these schools with a lists of high schools with a graduation rate above the city median (70 percent). They do not find that treated students, on average, apply for higher-performing schools but they do find that the treated students apply for schools where they are more likely to be admitted.[9]

We make three main contributions to this literature. Firstly, we can say more about the effects over the whole distribution of house-

[9]There has also been some semi-structural work using actual choice data, including previously mentioned studies by Ruijs and Oosterbeek (2017) and Hussain (2013).

holds, because our population is not restricted to (mainly) disadvantaged students in failing schools. Rather, it is a cross-section of the entire population in a rather well-off Swedish university town. This is important as it speaks to the potential effects school performance information can have on the competitive pressure on performance in the school system in general. Secondly, we can say more about why households react to the information. The school performance information we provided did not only include absolute results, (such as test scores and graduation rates) but also test scores adjusted for student composition. A school's average test scores are, in large part, determined by the school's student composition. Therefore, any changes in choice behavior driven by information on absolute performance could be due to parents looking for a stronger peer composition rather than a more productive school. This is however not the case for the adjusted test scores as, by construction, they are orthogonal to the student composition of the school.

We document a 5.2-percentage point (corresponding to 12 percent) increase in the probability of listing a top-performing school (defined as five highest performing schools) as one's top choice. The schools losing applications are the second-tier schools while the low-performing schools are unaffected. We also find that the effects on choice behavior are highly heterogeneous and mainly driven by high-skilled households and households with a native background. Further, these households mainly react to the measure of adjusted performance. This is, to our knowledge, the first documented effect of this kind. All in all, these findings indicate that school performance information interventions can have an effect on the competitive pressure to improve performance but that the effects might be local in the sense that low-performing schools will not be affected. Our findings are consistent with the findings in Hart and Figalo (2015) and Walters (2014). We do however fail to document the significant effects

of information on raw test scores on disadvantaged households choice behavior found in, for example, Hastings and Weinstein (2008) and Gallego et al. (2012).

Thirdly, we know how the placement mechanism in Linköping works and can replicate it. This combined with our design and our access to high-quality Swedish register data allows us to run simulations to investigate the potential effects of our treatment of a number of general equilibrium outcomes.[10] We find that the effects on choice behavior, at least in part, translate into effects on assignment with the share of students in the top-performing schools increasing with about 2 percentage points (this corresponds to about 40 percent of the increase in applications to these schools). This confirms the concerns expressed by Ruijs and Oosterbeek (2017) about looking at enrollment as a proxy for choice behavior in a setting where there might be capacity constraints. It is also in line with findings that the effects on enrollment of school quality information are larger in areas where capacity constraints do not bind (Hussain, 2013). This increase of students in the top-performing schools is mainly driven by high-skilled households and native households. These more advantaged households do, to an extent, displace students with a foreign background in the top-performing schools. All in all, this leads to an increased gap in school performance between low-skilled and high-skilled households and between foreign and native households. We also document a number of examples that illustrate other factors which one needs to take into account in order to predict how changes in choice patterns might affect assignment. These factors include ex-

[10] Again, due to data limitations, only a few previous studies have been able to analyze the effects of school performance information on other equilibrium outcomes than enrollment. A notable exception is a recent study by Abadie et al. (2017) that distributes report cards in a randomly selected set of villages in Pakistan. They find effects on total enrollment, average test scores and prices charged by private schools.

isting search patterns, the priority structure used in the school choice program as well as geographical factors.[11]

Finally, we look at school segregation. In doing so we speak directly to the second overarching question we posed in the beginning of this section: can school performance information interventions reduce school segregation? In a few previous studies there are results that, indirectly relate to the question of school segregation. Taken together these results are however inconclusive. For example, Hart and Figalo (2015) find that the share of high-SES students increases at highly-ranked schools when information on school rankings is released, while Mizala and Urquiola (2013), when conducting a similar exercise, find no significant effects. Corcoran et al. (2018) conduct subgroup analysis and find that disadvantaged households, in general do not seem to react more to school performance information compared to advantaged households while Hussain (2013) finds that poorer households seem to react more to simplified school performance information, when compared to other households.

We find suggestive evidence that school segregation in terms of parental skill-level increases somewhat as students from high-skilled households congregate in top-performing schools. We also document a significant decrease in school segregation in terms of migration background (6-points measured by the Duncan Dissimilarity Index). The reason for this is however not the one we might envisage: disadvantaged households with a foreign background becoming more prone to apply to schools with a larger share of students with a native background. Rather, as we have already mentioned, what we see is households with a native background applying to and, at least to an extent, being accepted into the top-performing schools. The top-performing schools in Linköping are, in general, schools with an above

[11] A priority structure is the rule used in school choice systems to determine which students are assigned to oversubscribed schools.

average share of students with a foreign background. This change accounts for about half of the observed reduction in segregation. When these students move into the top-performing schools they free up capacity at two oversubscribed middle-performing schools. These two schools have a low share of students with a foreign background and the seats freed up are, to a large extent, filled by students with a foreign background. This accounts for the other half of the reduction in segregation.

The remainder of this paper is organized as follows. Section 1.2 outlines the institutional setting and describes the experiment, Section 1.3 describes the data and the empirical strategy, Section 1.4 presents the results from the experiment, Section 1.5 discusses our simulation strategy and presents results on general equilibrium outcomes and Section 1.6 concludes.

1.2 The Institutional Setting and the Experiment

The Swedish compulsory school system consists of nine grades with children, in general, starting 1^{st} grade the year they turn seven. The Swedish school system is highly decentralized with the municipalities responsible for providing and financing compulsory education. School choice has been part of the Swedish school system since 1992. Our experiment was conducted in Linköping municipality in 2016.[12] In Section 1.2.1 we describe the school choice program in Linköping and the institutional framework surrounding it, and in Section 1.2.2 we describe the experiment we conducted.

[12]The reason for choosing Linköping municipality is that it was one of the few municipalities that was both interested in letting us conduct our experiment and could provide us with access to high-quality school choice data.

1.2.1 School Choice in Sweden and Linköping

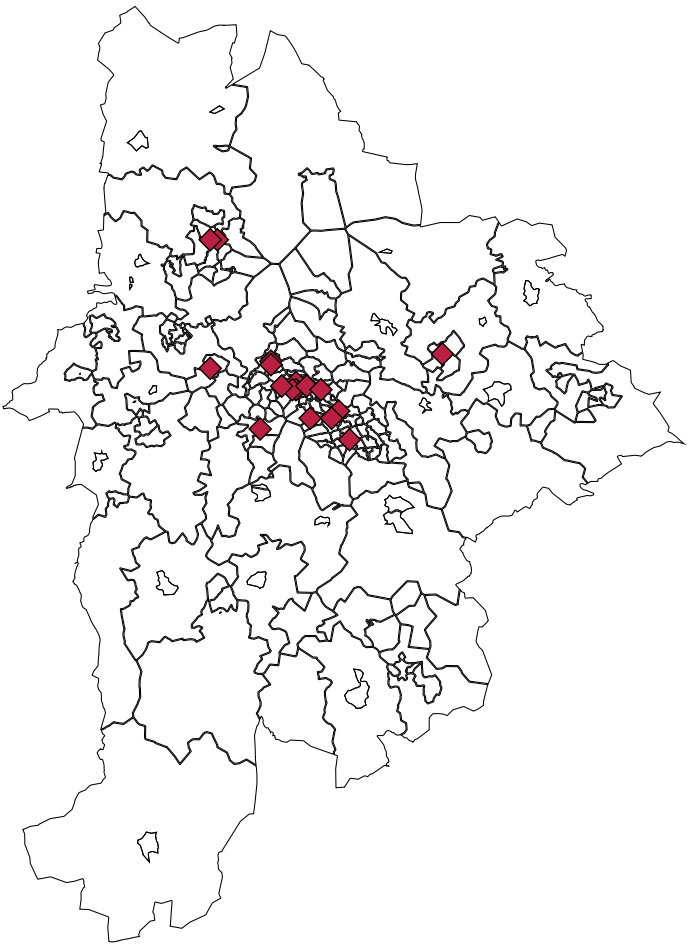
Linköping is the fifth largest municipality in Sweden with a total of 158,000 inhabitants. Most of them (109,000) live in the town of Linköping, located in the middle of the municipality and the rest live in smaller villages or rural areas surrounding the town. The average size of a student cohort is about 1,700. We focus on the choice of middle school (7^{th} to 9^{th} grade).[13] There are 20 middle schools in Linköping. 13 of them are public schools and seven are vouchers schools that are run by independent providers but publicly funded.[14] The middle schools in Linköping municipality are concentrated in the town of Linköping with the highest concentration just northwest of the city center. Figure 1.1 shows a map of Linköping municipality with the schools marked.

In the beginning of January of each year the municipality sends out a letter to all households with children about to start 7^{th} grade. The letter informs the parents that it is time to apply for middle school. The letter also includes information on what middle school is the household's default school. The default school is the school where the student will be placed if another school is not actively chosen or if they are not admitted to any of the schools they applied to. It is usually the school closest to the household's residence. In the letter there is also a user name and a password for an online tool, where

[13]The reason for focusing on the choice of middle school and not elementary school is that we think that school performance is relatively more important when choosing middle school compared to choosing elementary school. In Linköping the home to school distance seems to completely dominate the elementary school choice with 75 percent of all households opting for the closest school.

[14]In Linköping all the public schools are either elementary schools covering 1^{st} to 6^{th} grade or middle schools covering 7^{th} to 9^{th} grade. Five of the seven voucher schools that offer grades 7^{th} to 9^{th} in the municipality depart from this pattern by offering 4^{th} to 9^{th} grade or all grades from kindergarten or 1^{st} to 9^{th} grade. In total about 4 percent of students attend one of these five voucher schools in 6^{th} grade.

Figure 1.1: **Middle Schools in Linköping Municipality**



Note: The above map shows the location of the middle schools in in Linköping Municipality for which we have the location. The map also shows the division of the municipality into SAMS-areas (SAMS stands for Small Areas for Market Statistics and divides Sweden into about 9,500 small areas, based on municipal partitioning and electoral districts, depending on the size of the municipality. SAMS-areas are meant to have roughly the same population) in order to give a sense of population density.

the households can log in and submit their ranked preferences for the available schools.[15] Further, the letter contains information on how priority to oversubscribed schools is determined, where one can find more information on the available schools and information in other languages than Swedish.

The online tool is open for about three weeks at the end of January/beginning of February, during which parents can log in and submit their choices. They can list up to three ranked choices, and both public and voucher schools can be ranked. There are some middle schools in Linköping which offer special tracks such as drama, science, soccer and so on. If one wants to apply for one of these it has to be ranked as one of the three choices one is allowed. When the school choice period ends the municipality sends the lists of applicants to the voucher schools, so they can rank the applicants by their priority ordering. All voucher schools in Linköping use time spent in school-specific queues, sometimes combined with already enrolled siblings, to establish priority. The municipality also determines the priority each student has at the public schools they have chosen. Absolute distance from a students home to the school is used to determine priority unless one applies for a specific track for which there are special tests or tryouts.

When priority has been established the municipality places the students using a modified version of the Boston Mechanism.[16] In short, the municipality starts by giving everyone a seat at their default school. They then allow everyone to apply to their highest ranked school. Each school accepts students in order of priority up to their capacity limit. This is done in several rounds as once a student

[15]They can log in and accept the seat at their default school, this is however equivalent to not making a choice, as both actions (or lack thereof) would result in the child being placed at the default school.

[16]The Boston Mechanism is described in detail in Abdulkadiroglu and Sönmez (2003).

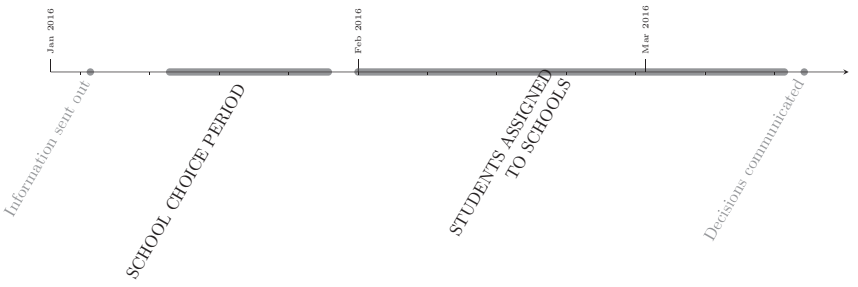
is accepted at a school other than their default school a seat at their default school is freed up, possibly allowing the municipality to give another student a seat at their highest ranked school. Once the municipality has placed everyone they can at their highest ranked school, they repeat this process for the students that did not get placed at highest ranked school and have listed a second choice. When that is done the process is repeated using the third choice (when it exists). The fact that the municipality uses the Boston Mechanism means that households have to take strategic incentives into consideration when choosing schools. This is because the mechanism drastically reduces the chances of being accepted at any particular school after being rejected from one's highest ranked school. In Linköping, only a few schools are, to any significant extent, chosen by students who do not have it as their default school and these schools are, in general, oversubscribed. This in combination with the use of the Boston Mechanism means that the second and third choices have a very marginal effect on the assignment of students.[17]

Once everyone is assigned to a school the municipality sends out a second letter informing the households which school their child has been assigned to. This usually happens in the middle/end of March. Figure 1.2 visualizes this process.

When making their choices the households have access to a number of different sources of information. The main source is the municipality's own online school comparison tool. This tool is available on the municipality's website and referred to in the first letter to the households. The tool covers both public and voucher schools in Linköping municipality. It contains information on the schools' location, the number of students in different grades, a short text written

[17] About 28 percent of students that do not get admitted to their most preferred school gets admitted to their second or third ranked option. This corresponds 3.2 percent of all students.

Figure 1.2: **Time-line of the school choice process**



by the schools themselves and results from a survey the municipality administers each year, where students and parents rate their school in terms of learning environment, facilities, food and so on. The tool does not contain any information on grades, test scores, graduation rates, student-teacher ratios or proportion of certified teachers. Information on these metrics can be found online in the Swedish National Agency for Education's database. The existence of this database is however not common knowledge and the database is not easy to navigate. The schools also offer an open house forum, so students and parents have a chance to visit the school before submitting their final choices. As there are 20 middle schools available in Linköping this is probably not an efficient way for parents and/or children to get a clear picture of their options. Given this, we think that it is fair to say that information on school performance in Linköping is limited.

1.2.2 The Experiment

The experiment took place in Linköping at the beginning of 2016. With the toss of a coin all households with a child expected to start 7^{th} grade were sorted into a treatment group and a control group. Simultaneously to receiving the letter from the municipality, house-

holds belonging to the treatment group also received a letter from us containing school performance information. For each available middle school in Linköping we listed the average performance on the national standardized tests. More specifically, for each school, we calculated the average test scores on the standardized test in 9^{th} grade in Swedish, mathematics and English over the last two years.[18] We also provided the households with a measure of how each school performs in relation to how it is expected to perform given its student composition.[19] We translated the letter into nine different languages to ensure that everyone would be able to understand it. We also included our contact details, should they have any questions. The letter in its entirety can be found in Appendix A.

1.3 Data and Empirical Strategy

In this section we describe the data and empirical strategy. In Section 1.3.1 we present our data sources, Section 1.3.2 presents some summary statistics and Section 1.3.3 outlines the empirical strategy and evaluates the identifying assumptions.

1.3.1 Data

There are three main sources of data used in this study. The first is the municipality of Linköping, the second is Statistics Sweden and the third is the Swedish National Agency of Education. The data from Linköping contains all school applications submitted in Linköping during the school choice period in 2016.[20] These applications are

[18] In Sweden standardized tests are taken in 3^{rd}, 6^{th} and 9^{th} grade.

[19] We will describe in detail how this measure is constructed in section 1.3.1.

[20] We also have access to choice data from 2014. This is not used in our main analysis but to test one of our identifying assumptions. We will discuss this in further detail in Section 1.3.3.

linked to household-level register data from Statistics Sweden, using individual identifiers and the Swedish multi-generational register.[21] From register data, we extract demographic and geographical information on all middle school starters as well as students already attending schools in Linköping. Further, for each household with at least one child about to start 7^{th} grade, we have an indicator for whether the household received our letter with additional information on school performance or not.[22] From the Swedish National Agency of Education, we have access to school performance (test scores and grades) and measures of student composition in all middle schools in Sweden. A more detailed description of the data will follow.

Definition of study population: We depart from the population of students that took part in the school choice process in Linköping in 2016. We drop two individuals where we do not observe school applications.[23] Further, we drop 41 individuals that were never assigned to the control or treatment group.[24] We also drop seven students enrolled in a religious school that, for reasons discussed later, will be dropped for our analysis. We end up with a population containing

[21]If the student is adopted, we use the characteristics of the adopting parents instead of the biological parents. We do this because we assume that the adopting parents are more likely to be involved in the school choice decision.

[22]In addition, we have an indicator for households that could not be located by the postal services where our letter was returned. This happened in three cases and all three were cases where the parents did not live at the same address. When parents did not live at the same address we sent out one letter to both addresses. Hence, at least one parent in all treated households received our letter.

[23]These individuals were not assigned a school by the municipality, making it likely that they moved out of Linköping during the school choice period and therefore did not take part in the school choice process.

[24]These individuals either moved to Linköping before the start of the school year, but after January 2016 when we defined our treatment and control group or they were students from other municipalities that, for some reason, took part in Linköping's school choice process.

1,608 households with children about to start 7^{th} grade, where we observe choices and relevant background characteristics and can categorize them as treated or not.

School choices: From the municipality, we obtained data on all school applications made during the school choice period in 2016. We observe up to three ranked choices for each child. We focus on top choices. We do this in order to keep the analysis clean and simple. As mentioned, the design of the system makes the second and third choice more or less inconsequential and a large fraction of households do not even bother to submit second or third choices. Given this, not much should be lost from this simplification. In addition to the applications we also observe each student's default school. If no school was actively chosen, we treat the default school as the top choice. Finally, we observe what school the student was assigned to by the municipality.

Demographic characteristics: From Statistics Sweden's LISA register, we have obtained information on the region of birth and level of education of all parents of all students in our population, as well as all students enrolled in a middle school in Linköping. We use this to create a variable indicating if the student has a foreign background (equal to one if the student is born abroad and/or both the parents are born abroad). We also create a variable indicating if the parents are high-skilled. This variable takes the value one if at least one of the parents have more than upper secondary education. We combine these indicators with the compulsory school register which links all students to their schools to construct measures of the student composition in the middle schools in Linköping.[25]

[25]The latest compulsory school register we have access to is from 2014 which means that we only observe students who were enrolled in 2014. Our measure

Geographical data: We observe the geographical location of both students and schools. Residential location of students is observed the first of January 2016. School location is observed in 2013.[26] Using this, we calculate the distance between each student and each school.[27]

School performance: Lastly, from the Swedish National Agency for Education we have access to the performance on the standardized tests in 9^{th} grade for all schools in Sweden. We use this data to construct the measures of school performance included in the letter sent out to the households in our treatment group. We observe the average scores (on a scale from 0 to 20, where 10 points is generally considered a passing score and 15 is considered a pass with distinction) for all tests in Swedish, mathematics and English. For each school, we calculate the average test score over all subjects. We further take the average over the years 2013-2014.[28] We use this as our measure of absolute performance. As mentioned we also provided the treated households with a measure on how the school performs in relation to how it is expected to perform given its student com-

of the student composition at the schools available to choose from in 2016 is therefore based on students enrolled in those schools in 2014. In general, we do not believe this is an issue because the student compositions of the schools do not vary much over time.

[26]We do not have access to school location later than 2013. We do not observe the exact location of schools or students but rather the midpoint of the 250m×250m grid where the student resides/the school is located. For those living in the rural areas of Linköping the grid is 1000m×1000m.

[27]The distance is calculated as $distance = \sqrt{(xcoord_{student} - xccord_{school})^2 + (ycoord_{student} - ycoord_{school})^2}$.

[28]We do this to reduce the noise that arises from the fact that certain schools have small student cohorts. See Kane and Staiger (2002) for a discussion on why this is important. We would like to have included additional years but due to data limitations this was not possible.

position. We calculate this by estimating the following regression on the full population of middle schools in Sweden, separately for 2013 and 2014:

$$Y_i = \alpha + \boldsymbol{X}_i'\boldsymbol{\beta}_x + \varepsilon_i \tag{1.1}$$

where Y_i is school i:s score on the standardized tests in 9^{th} grade and $\boldsymbol{X_i}$ is a vector of variables describing the student composition in school i.[29] We take the residuals from the schools in Linköping, average them over the years 2013-2014 and use that as a measure on how the school performs in relation to how it could be expected to perform given its student composition. Note that when we, in this paper, talk about the performance of a school we refer to the two above described measures. In doing so we are not claiming that these measures are in any way the best indicators of school quality. Rather, this is done to avoid carrying cumbersome definitions through the text.

We lack measures on performance and student composition for four of the 20 middle schools in Linköping. In two cases this is because the schools in question were opened after 2014 and therefore no tests had been administrated in these schools at the time we wrote our letter. One of these new schools is a subsidiary of an already existing school and the plan is for the new school to take over the role of this existing school. The new and the old school are located close to each other and students with either of these schools as default schools have them both as their default school. When our analysis demands it, we therefore categorize these two schools as the same school. The other new school is not linked to an existing school in

[29]The vector includes the share of boys in the school, the share of students with high-skilled parents and the share of newly arrived students. Our specification follows the Swedish National Agency for Education's method used in their so-called SALSA-measure.

the same way. When we need to rank and/or categorize this school we will assume that it is an average performing school. The third school is a small religious school (with between five and ten students in each grade). No one with this school as their default school, chooses another school and no one with another school as their default school chooses this school in either the control or the treatment group. We can therefore, without consequence to our analysis simply drop this school. The fourth school is another small school aimed at helping children from extremely rough conditions. No one has this school as their default school and we do not observe anyone applying to it. Hence, it can also be disregarded from in our analysis.

1.3.2 Descriptive Statistics

Table 1.1 reports the attributes of the available middle schools in Linköping. The schools average score is 13.32 (on a 20-grade scale). The average adjusted performance is -0.42. This means that, on average, the middle schools in Linköping are, given their student composition, underperforming somewhat compared to the average middle school in Sweden. On average, 22 percent of students in the schools have a foreign background and 62 percent of students have high-skilled parents. There is considerable variation between the schools both in terms of performance and student composition. For example, the average score at the lowest performing school is 9.75, which is not considered a passing grade while in the highest performing school the average score is 16.10 which is well over the threshold for a pass with distinction.

Table 1.2 reports descriptive statistics for the population of school choosing households. We can see that the demographic characteristics in the population are very similar to the student population in

Table 1.1: Characteristics of available schools

	Mean	SD	Min	Max
Performance				
Absolute performance	13.32	1.62	9.75	16.10
Adjusted performance	-0.42	0.79	-2.03	1.03
Student composition				
Foreign background (fraction)	0.22	0.21	0.03	0.86
High-skilled parents (fraction)	0.62	0.16	0.15	0.86
Observations	16			

Note: Absolute performance is the average score on the standardized tests in Swedish, mathematics and English in 9^{th} *grade during 2013-2014. These scores range from 0 to 20. Adjusted performance is the difference between the actual and expected score, where the expected score is the score predicted by the regression in Equation* (1.1)*. We classify a student as having a foreign background if a student is born abroad or if both parents are born abroad. We classify students as having high-skilled parents if at least one of their parents has an educational level higher than upper secondary.*

the average school. In general it is clear that our population is not as disadvantaged as in previous comparable studies (such as Hastings and Weinstein (2008) and Corcoran et al. (2018)). We can also see that the population's default schools, on average, are very similar to the schools in general. Further, we can see that households have, on average, six schools within a reasonable distance (within 5 kilometers) from their home. The variance in the number of schools close to home is large, with some households having no schools within 5 kilometers while others have more than ten. This reflects the fact that some households reside in the town of Linköping where the schools are concentrated while others reside in the smaller villages or the rural area surrounding the town. In most cases the schools close to home differ considerably in terms of performance and characteristics. The average difference in test scores between the highest performing and the lowest performing school close to home (5km) is over 4 points. Just like with the number of schools close to home there is a lot of variation in these measures.

This variation in school characteristics of available schools in combination with the fact that most schools are located in the city center (see Figure 1.1) where the municipality's bus lines connect, implies that most households have the opportunity to react to our treatment by choosing higher performing schools.

Finally, in Table 1.3 we describe the choice behavior in the control group. In Panel A we can see that about 50 percent of the households choose a school other than their default school. In general households with a native background and high-skilled households are more prone to opt out of their default schools compared to households with a foreign background and low-skilled households. Native and high-skilled households also choose schools with better absolute performance, a larger fraction of students with high-skilled parents and smaller share of students with a foreign background compared

Table 1.2: Background characteristics of households

	Mean	SD	Min	Max
Predetermined variables				
Foreign background	0.22	0.41	0.00	1.00
High-skilled parents	0.66	0.47	0.00	1.00
Characteristics of default schools				
Abs. performance	13.16	1.43	9.75	16.10
Adj. performance	-0.41	0.65	-2.03	1.03
Share foreign	0.23	0.19	0.03	0.86
Share high-skilled	0.60	0.14	0.15	0.86
Within 5000 meters from home				
No. of schools	6.00	4.74	0.00	13.00
Diff (max-min), abs. performance	4.05	2.30	0.00	6.35
Diff (max-min), adj. performance	2.12	1.09	0.00	3.06
Diff (min-max), share foreign	0.46	0.32	0.00	0.81
Diff (min-max), share high-skilled	0.43	0.25	0.00	0.71
Observations	1608			

Note: The absolute performance is the average score on the standardized tests in Swedish, mathematics and English in 9^{th} *grade during 2013-2014. These scores range from 0 to 20. The adjusted performance is the difference between the actual and expected score, where the expected score is the score predicted by the regression in Equation* (1.1)*. We classify a student as having a foreign background if a student is born abroad or if both parents are born abroad. We classify students as having high-skilled parents if at least one of their parents has an educational level higher than upper secondary.*

to households with a foreign background and low-skilled households. When it comes to adjusted performance, this pattern is not as clear. High-skilled households choose higher performing schools compared to low-skilled households, but households with a foreign background choose higher performing schools compared to households with a native background.

In Panel B we report the difference between the top-ranked school and the default school in various school attributes for the subset of households that opt out of their default schools. Conditional on opting out, all groups, on average, choose schools with better absolute and adjusted performance compared to their default school. All groups do also, on average, choose schools with a larger fraction of students with high-skilled parents and a lower fraction of students with a foreign background compared to their default schools. When looking at adjusted performance, we see that high-skilled households and native households improve more compared to low-skilled households and households with a foreign background. Households with a foreign background and low-skilled households also choose schools with a smaller share of students with foreign background compared to their default schools than the other groups.

All in all this is in line with the previous findings that more disadvantaged households seem to choose lower performing schools compared to more advantaged households (as in Hastings et al. (2009), Abdulkadiroglu et al. (2017) and Abdulkadiroglu et al. (2018)).

1.3.3 Empirical Specification

Estimating treatment effects in a randomized controlled trial is straightforward. Let T_i be an indicator equal to one if household i is assigned to the treatment group and zero otherwise. The effect of the treatment on an outcome Y_i can then be estimated using a linear regression model:

Table 1.3: Choice behavior in the control group

	All	Native	Foreign	High-skilled	Low-skilled
	A. Whole sample (top)				
Choosing non-default	0.50	0.56	0.24	0.56	0.39
Distance	4345	4840	2690	4439	4240
Abs. performance	13.81	13.91	13.47	14.05	13.38
Adj. performance	-0.21	-0.23	-0.13	-0.17	-0.29
Share foreign	0.21	0.16	0.36	0.18	0.24
Shar high-skilled	0.67	0.69	0.61	0.70	0.62
Observations	783	615	166	510	265
	B. Conditional on opting out (top-default)				
Distance	1956	2047	1296	1931	2037
Abs. performance	0.90	0.92	0.78	0.89	0.92
Adj. performance	0.49	0.53	0.22	0.52	0.40
Share foreign	-0.07	-0.06	-0.13	-0.06	-0.11
Share high-skilled	0.17	0.17	0.15	0.16	0.18
Observations	318	280	37	244	74

Note: Absolute performance is the average score on the standardized tests in Swedish, mathematics and English 9^{th} *grade during 2013-2014. These scores range from 0 to 20. Adjusted performance is the difference between the actual and expected score, where the expected score is the score predicted by the regression in Equation* (1.1)*. We classify a student as having a foreign background if a student is born abroad or if both parents are born abroad. We classify students as having high-skilled parents if at least one of their parents has an educational level higher than upper secondary.*

$$Y_i = \alpha + \beta T_i + \epsilon_i \tag{1.2}$$

where ϵ_i is a (heteroskedasticity robust) error term. We are interested in β, the effect of our treatment on outcome Y.[30]

For β to have a causal interpretation, we need the randomization of households into the control and the treatment group to have been successful. To confirm random assignment, we run a regression with the treatment indicator (T_i) as the dependent variable and include all variables (from Table 1.2) as independent variables. If randomization was successful, we do not expect these variables to predict treatment. Table 1.4 shows the results from this regression, as well as the p-value from an F-test of joint significance. One out of eleven of the included predetermined variables is significant (at the 10%-level), and the F-test confirms the null hypothesis. We conclude that randomization was successful.

Another potential threat to identification is that information could have spread from treated households to households in the control group. There is no perfect test of this and hence it is not possible to exclude this possibility. However, the time window for school choice it not that long (less than three weeks), limiting the scope for information to spread. During this time period, we covered all local media and there was no mention of the experiment. Anecdotal evidence from households contacting us also suggest that the information sent out was not subject to much discussion between parents.[31] Further,

[30] As neither sampling nor assignment of treatment was clustered, we choose not to report clustered standard errors. See Abadie et al. (2017) for a discussion about when to use clustered standard errors. All statistical tests in this paper have also been run using randomization inference (see Young (2017) for more details on randomization inference). Doing this does not affect the results.

[31] Of 30 individuals contacting us regarding the information sent out, one had discussed it with another parent and one had heard another parent mentioning the letter.

Table 1.4: Test of randomization

	Treated
Foreign background	0.0317
	(0.0381)
High-skilled	-0.0101
	(0.0319)
Abs. performance (default school)	0.0440
	(0.0452)
Adj. performance (default school)	-0.0714
	(0.0513)
Share foreign (default school)	0.0529
	(0.137)
Share high-skilled (default school)	-0.187
	(0.383)
No. of schools within 5000m	-0.00853
	(0.0113)
Diff (max-min), abs. performance (schools within 5000m)	-0.0307
	(0.0437)
Diff (max-min), adj. performance (schools within 5000m)	0.0777*
	(0.0471)
Diff (min-max), share foreign (school within 5000m)	-0.105
	(0.241)
Diff (min-max), share high-skilled (schools within 5000m)	0.286
	(0.227)
Observations	1608
P-value for F-test of joint significance	0.601

Note: Standard errors in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Dependent variable is the treatment indicator (T_i). Absolute performance is the average score on the standardized tests in Swedish, mathematics and English in* 9^{th} *grade during 2013-2014. Adjusted performance is the difference between the actual and expected score, where the expected score is the score predicted by the regression in Equation* (1.1). *We classify a student as having a foreign background if a student is born abroad or if both parents are born abroad. We classify students as having high-skilled parents if at least one of their parents has an educational level higher than upper secondary.*

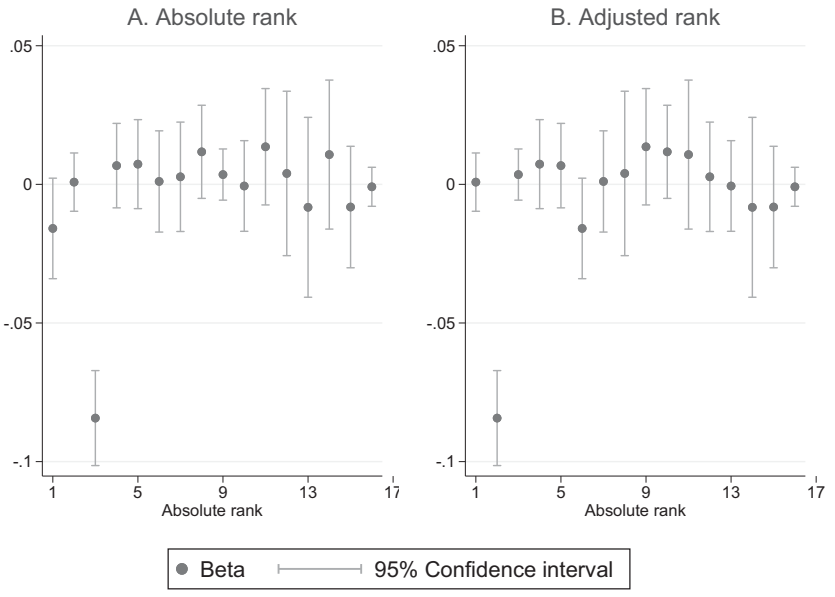
we have access to the choices made by households with children about to start 7^{th} grade in Linköping in 2014. If information did not spread from the treatment group to the control group the school choosing households in 2014 should have access to the same information as our control group in 2016 and hence make similar choices.[32] For the 16 schools that were available to choose from in both 2014 and 2016 we estimate the differences in probability of the school being chosen in 2014 and 2016 by running the following regression:

$$S_i = \alpha + \beta 2016_i + \epsilon_i \tag{1.3}$$

where S_i is an indicator variable taking the value of one if the school in question was household i:s top-ranked school and zero otherwise, 2016_i is an indicator variable if the household i is part of our control group in 2016 and zero otherwise and ϵ_i is a (heteroskedasticity robust) error term. β is then the difference in the total fraction of students choosing the school in question in the control group in 2016 and in the general population in 2014. In Figure 1.3 we plot these β:s and the related confidence intervals for each school. We order the schools by the performance measures we included in our letter. If the information has spread from the treatment group to the control group we would expect the β:s to increase with performance. We only have one significant coefficient: one of the low performing schools is considerably less chosen in the control group of our experiment compared to the general population of households 2014. The coefficient is very large, the fraction of students choosing the school is reduced by almost 8 percentage points. The school went from over a hundred applicants in 2014 to only eleven in 2016. We know that one middle school in Linköping was evacuated in 2015 due to extremely poor working conditions. The school was temporarily housed in the

[32]Time-trends could affect these results and hence this is not a perfect test.

Figure 1.3: Test of information dispersion



Note: The above figure shows the β:s for each school estimated in regression (1.3) with schools sorted by their absolute and adjusted performance respectively.

cellar of a nearby high school and local media reported that students and teachers alike were fleeing the school. We think that this, rather than our treatment, explains this extreme outlier. Except for this single school we do not see an upwards slope (or indeed any other pattern that would suggest that information had spread between between the treatment group and the control group). Given this we are comfortable with assuming that the spread of information from the treatment group to the control group was limited.

1.4 Results from the Experiment

In Table 1.5 we report the treatment effects on the share of households choosing to rank another school than their default school as their top choice. This can be thought of as an effect on the extensive margin. We also report effects on the composition of those ranking another school than their default school as their top choice. In column 1, we see that there is no significant aggregate effect on the share of households opting out of their default school. The same is true for the composition of households opting out of their default school (column 2 and 3). Hence, in this setting, it does not seem like providing easily available information on school performance induces additional (or different) households to try to opt out of their default schools.

Table 1.5: Effects on the extensive margin

	Active	Share of active households that are high-skilled	Share of active households with foreign background
Treated	0.0109	0.0251	0.0314
	(0.0250)	(0.0307)	(0.0227)
Constant	0.499***	0.732***	0.103***
	(0.0179)	(0.0225)	(0.0155)
Observations	1608	812	810

*Note: Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We classify a student as having a foreign background if a student is born abroad or if both parents are born abroad. We classify students as having high-skilled parents if at least one of their parents has an educational level higher than upper secondary. Active choice is an indicator variable taking value one if the households opts out of their default school and zero otherwise.*

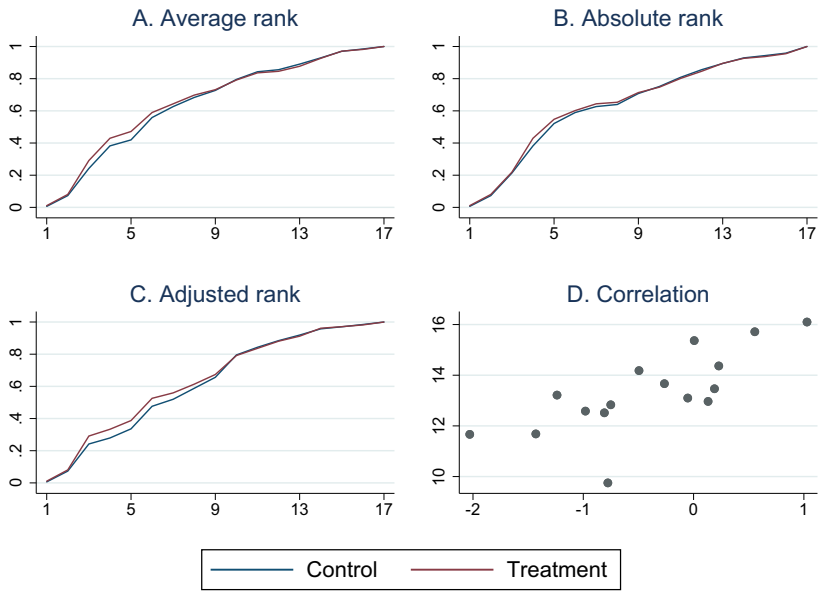
We move on to the performance (defined by the measures of per-

formance we included in our letter) of the top-ranked schools. Figure 1.4 shows the cumulative distribution of top choices over the performance rank of the schools in the control group and treatment group respectively. In Subfigure A the schools are ranked in terms of their average performance according to the two performance measures that were included in our letter.[33] We can see that the treatment group is more likely to choose the top-performing schools with the gap between the groups widening up until the fifth highest-performing school. The gap between the distributions then starts to close and when we reach the school with the average rank of ten the gap is completely gone. Hence, it seems like our treatment shifts applications from the second-tier schools to the top-performing schools. In Subfigure B and C we show the cumulative distribution of top choices over the absolute rank and adjusted rank of the school in the control group and treatment group respectively. We can see that, in both figures, the pattern from Subfigure A remains with the gap between the groups initially widening and then closing. This should not be surprising given the high correlation between the two measures (as seen in Subfigure D). The size of the gap is however considerably smaller when looking at the rank in terms of absolute performance compared to when looking at the rank in terms of adjusted performance.

Table 1.6 confirms and quantifies these results. Panel A corresponds to Subfigure A in Figure 1.4 with the schools ranked by average performance. Applications to top-performing schools increase with about 5.2 percentage points (corresponding to 12 percent). Applications to mid-tier schools decrease by 5.6 percentage points (corresponding to 15 percent). In Panel B the schools are ranked by absolute performance and in Panel C by adjusted performance. When ranking the schools by adjusted performance we can see that the coefficients and significance levels are almost identical to those we

[33](Absolute rank+Adjusted rank)/2.

Figure 1.4: Cumulative distribution of top-ranked schools over school rank



Note: Subfigure A-C shows the cumulative distribution of top-ranked schools over the average, absolute and adjusted performance of the schools (defined by the measures included in our letter) for the control group and the treatment group respectively. Subfigure D shows the correlation between the adjusted performance (x-axis) and the absolute performance (y-axis).

get when ranking the schools by average performance (in Panel A). When ranking the schools over absolute performance the effects are no longer significant and the coefficients are smaller.

Table 1.6: Effects on performance of top-ranked schools

	Top-performing	Middle-performing	Low-performing
A: Schools ranked by average performance			
Treated	0.0522**	-0.0557**	0.00355
	(0.0248)	(0.0238)	(0.0202)
Constant	0.420***	0.376***	0.205***
	(0.0177)	(0.0174)	(0.0145)
B: Schools ranked by absolute performance			
Treated	0.0311	-0.0306	0.00432
	(0.0247)	(0.0205)	(0.0216)
Constant	0.559***	0.230***	0.248***
	(0.0178)	(0.0151)	(0.0155)
C: Schools ranked by adjusted performance			
Treated	0.0516**	-0.0551**	0.00355
	(0.0239)	(0.0247)	(0.0202)
Constant	0.336***	0.459***	0.205***
	(0.0170)	(0.0179)	(0.0145)
Observations	1608	1608	1608

*Note: Standard errors in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variables Top-performing, Middle-performing and Low-performing are indicator variables that take the value one if a household's top-ranked school is ranked (average) 1-5, 6-10 or 11+ and zero otherwise.*

All in all, looking at the distributions in Figure 1.4 it seems like the average rank of the schools is the best metric to use when evalu-

ating the effects on our treatment. The results from Subfigure B and C combined with the results in Table 1.6 indicate that if one dimension of our information was more important it was the information on adjusted performance of the schools. This could be because households have already taken the absolute performance into account when making their choices, either because they actually have access to the information on absolute performance directly or because they have a good enough proxy (such as student composition or reputation). It could also be because they are aware of the fact that the absolute performance of a school is not necessarily a good measure of school quality. These results indicate that the changes in choice behavior we observe are not driven by households looking for a stronger (in terms of observable characteristics) peer group, but rather something else.

To further understand this shift in school choice behavior, we investigate whether the effect is driven by specific groups of households. Table 1.7 reports results for subgroups. In Panel A we focus on high-skilled households. Among this group, applications to top-performing schools increase with about 8.1 percentage points. Applications to middle-performing schools decrease by 6.7 percentage points. In Panel B, where we report results for low-skilled households, there are no significant effects. A test for the different treatment effects between the two groups is close to significant (p-value = 0.116). In Panel C and D we instead cut the sample by their migration background. Among native households applications to top-performing schools increase with about 7.6 percentage points. Applications to middle-performing schools decrease by 6.8 percentage points. We see no significant effects for households with foreign background. The difference in treatment effects between the groups is significant (p-value = 0.073).

The above documented shift in applications from second-tier to top-performing schools could potentially increase the competitive

Table 1.7: Effects on performance of top-ranked schools by household type

	Top-performing	Middle-performing	Low-performing
	A: High-skilled		
Treated	0.0808***	-0.0669**	-0.0139
	(0.0308)	(0.0289)	(0.0225)
Constant	0.476***	0.360***	0.163***
	(0.0222)	(0.0213)	(0.0164)
Observations	1050	1050	1050
	B: Low-skilled		
Treated	0.00143	-0.0447	0.0433
	(0.0400)	(0.0419)	(0.0393)
Constant	0.314***	0.410***	0.276***
	(0.0288)	(0.0305)	(0.0277)
Observations	543	543	543
	C: Native		
Treated	0.0758***	-0.0685**	-0.00730
	(0.0279)	(0.0272)	(0.0220)
Constant	0.403***	0.407***	0.190***
	(0.0198)	(0.0198)	(0.0158)
Observations	1260	1260	1260
	D: Foreign		
Treated	-0.0332	-0.00567	0.0388
	(0.0541)	(0.0476)	(0.0484)
Constant	0.481***	0.263***	0.256***
	(0.0396)	(0.0349)	(0.0346)
Observations	343	343	343

*Note: Standard errors in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. See Table 1.6 for definitions of the dependent variables and Table 1.5 for definitions of foreign background and high-skilled.*

pressure on schools to improve their performance. Note also that the strategic incentives present in the school choice program in Linköping might put a downward pressure on the effects on choice behavior that we observe. In that sense, the effects should be interpreted as a lower bound for the information effect in a strategy proof school choice program. Finally, the effects suggest a substitution from middle- to top-performing schools driven mainly by native and high-skilled households. This may imply that this increased competitive pressure would potentially not benefit everyone the same way and that disadvantaged students in low-performing schools can get left behind.

In general, this type of effect, where information on a more sophisticated type of performance measure makes high-skilled and native households switch their top-choices from the second-tier schools to the top-performing schools has, to our knowledge, not previously been documented. They are however in line with the findings in Hart and Figalo (2015) and Walters (2014). The fact that we do not find any significant effects among more disadvantaged households of information on test scores is in contrast to Hastings and Weinstein (2008) and Gallego et al. (2012). There are a number of potential reasons for this, one is that Sweden is a more equal society than Chile or the US and thereby differences between schools are not as large and disadvantaged households might not be as uninformed.[34] It could also be because we couple the information on absolute performance with information on adjusted performance which they do not. This could potentially make the information less transparent and more confusing for disadvantaged households (as discussed in Imberman and Lovenheim (2016)).

[34]For example in Hastings and Weinstein (2008) only 11 percent of students select out of their default school before receiving information, in our sample the corresponding number is 24 percent.

1.5 General Equilibrium Effects

This section is organized as follows. Section 1.5.1 describes our simulation strategy, Section 1.5.2 reports the effects on assignment and finally, Section 1.5.3 investigates the effects on segregation.

1.5.1 Simulation

In order to look at general equilibrium outcomes we run a simulation where we replicate the assignment mechanism and the priority structure used in Linköping. We could compare assignment in the treatment and the control group directly but as the assignment of one student can be dependent on the assignment of another student we risk overestimating the effects on assignment by doing so as only half of the households have access to the additional performance information included in our letter.

We continue to focus on the households' top-choice in order to keep our analysis consistent. Further, we are only looking at the direct (or short-term) effects on student assignment. For example, it is possible that the effects on student assignment might trigger reactions from teachers and school choosers in coming years. These types of dynamic effects are not taken into account in this simulation.

What we do is the following:

1. We approximate each school's capacity with either the number of default seats at the school or the number of students assigned to the school, whichever is higher.[35]

2. We draw, with replacement, from the control group, a cohort of the same size as our observed cohort.

[35]Note that this might mean that we underestimate the capacity of the schools that are not filled to capacity. Given that these are the less popular schools we do not think this will be a concern.

3. We place all students at their default school (as observed in the data).
4. We allow everyone to apply for their top-ranked school (again, as observed in the data).
5. We accept students applying to schools where there is free capacity in order of their priority.[36] Students not accepted at their preferred school are referred back to their default school.
6. We document the number of students leaving each school and readjust the free capacity accordingly.
7. We repeat step 4-6 until no students are placed in new schools.
8. We document relevant measures of the assignment.
9. We repeat step 2-8 one thousand times.
10. We repeat step 2-9 for the treatment group.

The difference in the recorded metrics from the simulations where we drew our sample from the control group with the simulations where we drew our sample from the treatment group, can then be interpreted as the effect of the treatment. When it comes to the metrics of interest we record the share of students admitted to the top-, middle- and low-performing schools. We also record the share of students in different subgroups in the top-performing schools. Further, we document the average rank of the assigned school for the whole population and the subgroups.

We also look at segregation. Segregation between groups divided into organizational units (such as schools) can be measured using

[36]We use the distance between the students home and the school to establish priority for public schools while, for voucher schools we use a random number. If a school is a student's default school, we give them top priority.

either measures of evenness or measures of exposure (Massey and Denton, 1988). Exposure measures are sensitive to the share of minority students in the population, something that evenness measures are generally not. As Allen and Vignoles (2007) point out the share of minority students is not something that educational policy can directly affect and therefore we use a measure of evenness when evaluating the effects of information on school performance on school segregation.

We follow Massey and Denton (1988) and report the Duncan Dissimilarity Index. The Dissimilarity Index (DDI) is given by:

$$DDI = \frac{1}{2} \sum_{j=1}^{N} \mid \frac{a_j}{a} - \frac{b_j}{b} \mid$$

where N is the number of schools, a_j is the number of students from group A in school j, a is the total number of students from group A, b_j is the number of students from group B in school j and b is the total number of students from group B. The Dissimilarity Index ranges from zero (no segregation) to one (total segregation) and has a clear interpretation as the percentage of one of the two groups that would have to move to a different organizational unit (school), in order to produce a distribution in each organizational unit that matches the distribution of the entire population (zero segregation) (Duncan and Duncan, 1955).

The Duncan Dissimilarity index is however not without it's drawbacks. Most notably, it fails the transfer principal (James and Taeuber (1985) and White (1986)). This means that it is insensitive to the redistribution of minority group members among organizational units with minority proportions above or below the overall minority proportion. Only transfers of minority members from units where they are overrepresented to units where they are underrepresented (or vice versa) affect the value of the index. Therefore, we will also

report the effects on the Thiel Entorpy Index. This index was originally proposed by Theil and Finizza (1971)) and is another measure of evenness for which the transfer principle holds. The Thiel Entorpy Index (TEI) is given by:

$$TEI = \frac{1}{N} \sum_{i=1}^{N} \frac{x_j}{\mu} ln(\frac{x_j}{\mu})$$

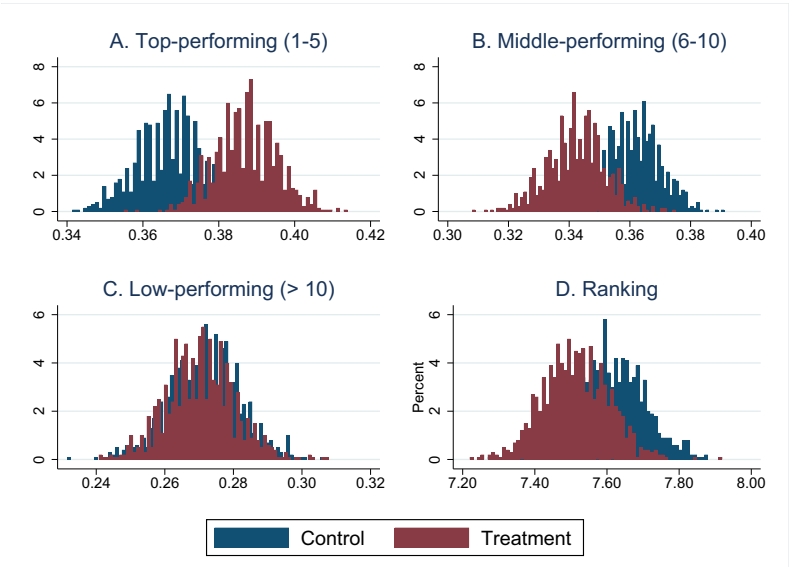
where N is the number of schools, x_j is a characteristic of school j (such as the share of students with a foreign background assigned to school j) and μ is the mean of x for all schools in N. The Thiel Entropy Index can be interpreted as the difference between the diversity (entropy) of the system and the weighted average diversity of individual organizational units, expressed as a fraction of the total diversity of the system (Reardon and Firebaugh, 2002).

1.5.2 Effects on Assignment

Figure 1.5 shows the simulation results for the share of students that are admitted to top-, middle- and low-performing schools for the control and treatment group separately. In Subfigure A we see that the share assigned to top-performing schools increases with 2.1 percentage points from our treatment (from 36.6 percent when we sample from the control the group to 38.7 percent when we sample from the treatment group). This corresponds to 40 percent of the effects we documented on choice behavior in Table 1.6. This implies that in a setting where no one has access to information on school performance there would be free capacity in the top-performing schools while in a setting where this information was available to everyone the top performing schools would be oversubscribed. Therefore, the effect of our treatment on assignment of those that react to the treatment is muted by institutional factors, in this case capacity constraints. We

also see, in Subfigure B, a drop of 1.9 percentage points in the share of students assigned to the middle-performing schools (from 36.1 percent to 34.2 percent). In Subfigure C we can see that assignment to the low-performing schools is not affected much. Finally, in Subfigure D, we can see that this is reflected in the average rank of the schools attended with students, on average, attending school with a 0.1 better ranking.

Figure 1.5: Share attending different types of schools



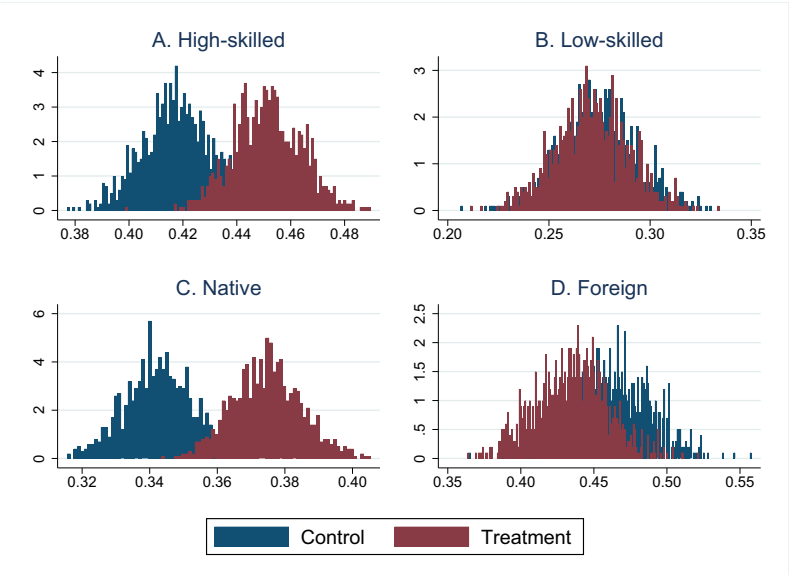
In Figure 1.6 we report the effects on being assigned to a top-performing school for the different subgroups. In Subfigure A we see that the share of students with high-skilled parents assigned top-performing schools increases by 3.3 percentage points (from 41.8 percent to 45.1 percent). In Subfigure C we see a similar increase for students with a native background (3.4 percentage points increase from 34.2 percent to 37.6 percent). This is what we should expect

as it was in these groups of households where we documented significant effects on choice behavior. Subfigure B shows that the share of students with low-skilled parents in the top-performing schools is unaffected. Subfigure D shows that the share of students with foreign background in the top-performing schools decreases by 2.3 percentage points (from 46.1 percent to 43.8 percent). All students with foreign background with a top-performing school as their default school in the treatment group are assigned to a top-performing school. This implies that the decrease in the share of students with a foreign background assigned to a top-performing school is due to them being displaced by the increased pressure from other students applying to these schools.[37]

Figure 1.7 shows the simulation results for the average ranking of the assigned school for different subgroups. In Subfigure A and C we can see that students of high-skilled parents and native students average ranking is reduced, meaning that they on average, get into better schools. Subfigure B reports the results for students with low-skilled parents. We can see that the average ranking is not affected. This is expected as we did not see any change in their choice behavior or any evidence of them being displaced from the top-performing school by high-skilled households. Subfigure D reports the results for students with a foreign background. We can see that the average ranking of their assigned schools increases, implying that they on average, are assigned to lower performing schools. All in all, the additional performance information in our letter increases the difference

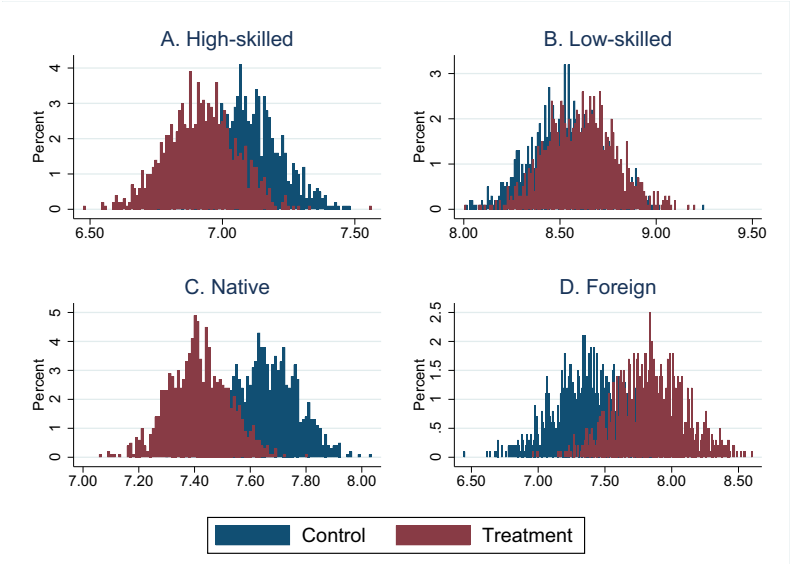
[37]In untabulated results we have also examined if the effects vary with the performance of the households' default school or the households' distance to the top-performing schools. Given the design of the school choice system in Linköping it seems likely that this could be the case. We do find that the treatment effect is mainly driven by households with high-performing default schools. We can however not distinguish if the reason for this is selection, the design of the system or something else.

Figure 1.6: Share attending a top-performing schools by household type



in school performance between students from high-skilled and low-skilled households and revereses the school performance gap between students with a native and foreign background.

Figure 1.7: Average school ranking, by household type



Case Study: The Schools Loosing Applicants

The above establishes that the observed changes in choice behavior translate into changes in assignment. Our design allows us to dig a little deeper into the relationship between changes in choice behavior and changes in allocation of students. Given the limited number of schools in Linköping this exercise will take the form of a case study. The findings in this subsection should therefore be read as such.

Specifically, we will look at the middle-performing schools that loose applicants and investigate how this loss in applicants translates

into loss in enrollment. In Subfigure A in Figure 1.8 we plot the loss in applications and the loss of enrollment for these five schools. We can see that one school is more or less unaffected in both dimensions. Four schools do however experience considerable drops in terms of applications. Among these, three do not experience the same sort of reduction in enrollment, while one does.

One reason for why a school's enrollment might not drop when loosing applicants is that is was oversubscribed before the shock and so can retain its enrollment rate even when loosing applicants. Subfigure B in Figure 1.8 shows the relationship between the change in application and the school's popularity (measured by the subscription rate). Two of the schools are oversubscribed. These two schools are also schools experiencing considerable drops in terms of applications but not in terms of enrollment. For these two schools the loss of applicants does not translate into a loss in enrollment as they were already oversubscribed.

Subfigure B in Figure 1.8 does however indicate that the pre-shock subscription rate does not explain everything. There are two schools that both experienced considerable drops in applications, have about the same pre-shock subscription rate but where one experiences a considerable drop in enrollment while the other one does not. Subfigure C in Figure 1.8 shows the distribution of the distance of the students with either of these two schools as their default school and the schools that are either oversubscribed and/or top-performing. The students that are allocated to the school that looses enrollment, live much closer to the oversubscribed and/or top-performing schools compared to the students who are allocated the other school. As priority is based on distance the students from the first (closer) school will be admitted to these schools before the students in the second school. Any free capacity in these schools will therefore first be filled with students from the first (closer) school. Students in the sec-

ond school simply cannot get in to the oversubscribed and/or top-performing schools and so it does not matter if they become more prone to choosing these schools.

All in all, this implies that when trying to translate changes in choice behavior into changes in enrollment we need to take a number of factors into account including the already existing choice behavior, the allocation mechanism and priority structure and the geography.

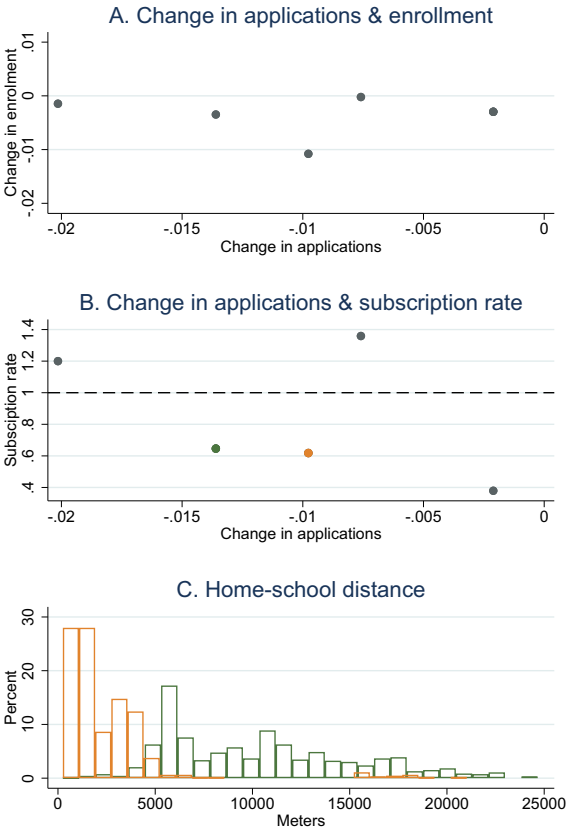
1.5.3 Effects on Segregation

Figure 1.9 shows the effects on segregation. We start by looking at school segregation in terms of parental skill-level. Subfigure A and B report simulation results for segregation in terms of parental skill level using the Duncan Dissimilarity Index and the Theil Entropy Index respectively. There is no clear effect on the Duncan Dissimilarity Index but there is a 2-point increase in the Theil Entropy Index. Segregation in terms of parental skill-level seems to increase somewhat in the top of the distribution as even more students from high-skilled households congregate in the top-performing schools.[38]

Subfigure C and D report simulation results for segregation in terms of migration background. In Subfigure C we can see that the Duncan Dissimilarity Index is considerably lower when we sample from the treatment group compared to when we sample from the control group. It drops a total of 6 points from 0.44 to 0.38 (corresponding to 14 percent). Subfigure D shows a similar drop in the Theil Entropy Index. This is a considerable drop in school segregation and, in fact, segregation in a setting with school choice where everyone had access to our school performance information would

[38]The top-performing schools were before treatment the schools with the highest share of students with high-skilled parents (71.4 percent compared to 66 percent in the general population).

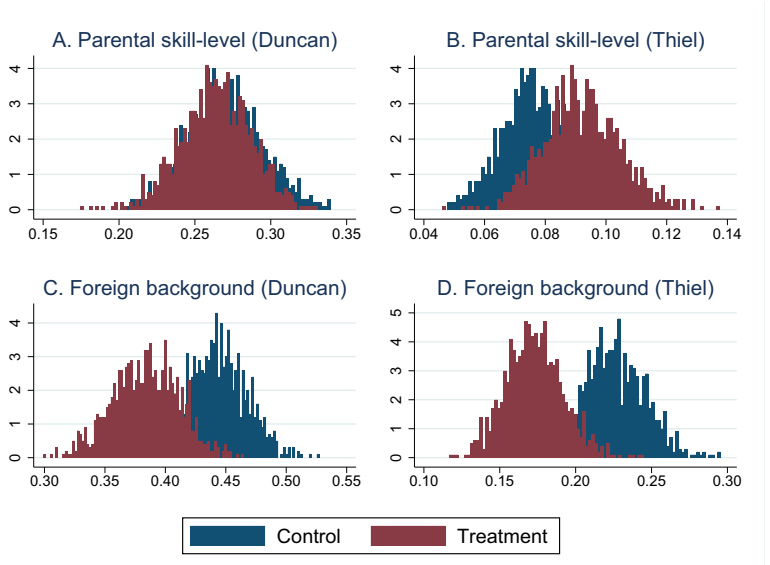
Figure 1.8: Case study on the middle-performing schools



Note: Panel A shows the change in applications and enrollment for the middle-performing schools. Panel B shows the relationship between the change in applications and the subscription rate of the middle-performing schools. Panel C shows distribution of distance from the home of the students with one of the two schools marked in Panel B as their default to oversubscribed and/or top-performing schools.

be lower than it would be if everyone stayed in their default school (DDI=0.41).[39]

Figure 1.9: Effects on segregation



The top-performing schools have, on average, a higher share of students with foreign background compared to the second-tier schools (23 percent and 13 percent respectively). Given the interpretation of

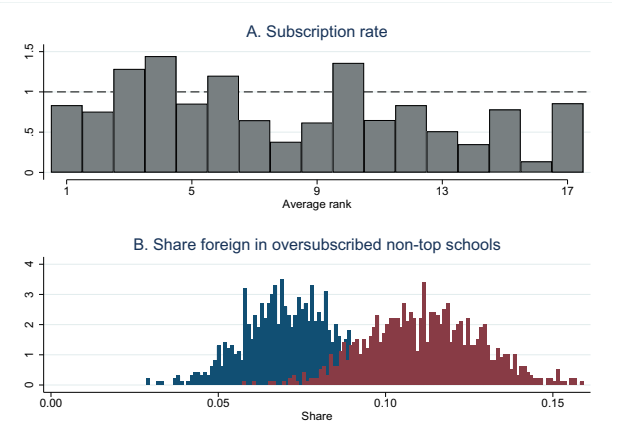
[39]Carrington and Troske (1997) show that evenness measures are sensitive to randomness when the average size of the organizational unit is small. As we are interested in the difference in unevenness between two scenarios where the size of the organizational units is more or less the same we do not think this should be a big problem. As a safety measure we will however simulate a random assignment of students from the treatment group and the control group separately and compare the distribution of the *DDI* in these two cases to make sure that there is no underlying difference due to randomness. We also include a measure of exposure, Isolation (See Lieberson (1981) for a description) to confirm that our results are not driven by this property of the evenness measures. Results from these robustness checks can be found in Appendix B.

the Duncan Dissimilarity Index the 3.4 percentage point increase of native students assigned to the top-performing schools can therefore explain, at most, just over half of observed effect on segregation.

Given that we did not see any other significant effects on choice behavior it seems likely that the the rest of the effect is from the propagation of the original effect through the system. Subfigure A in Figure 1.10 shows the schools subscription rate with the schools ordered after average performance. We can see that there are two oversubscribed schools outside the top-performing school (schools ranked 1-5). We know from Figure 1.8 that these two schools experienced a significant reduction in terms of applications but not in terms of enrollment. In Subfigure B in Figure 1.10 we report how the share of students with a foreign background assigned to these two schools changes with our treatment. We see a significant increase of 3.8 percentage points (from 7.4 percent to 11.2 percent). This effect combined with the increased share of native students in the group of top-performing school corresponds well to the total observed decrease in the Duncan Dissimilarity Index.

This reduction in segregation is not the one usually envisioned as it is not driven by students with a foreign background applying and getting into schools with a lower share of students with a foreign background. Rather it is driven by native students applying and getting into schools with a higher share of students with foreign background (as these are the highest performing schools). In doing so, they free up space in schools with a low share of student with a foreign background. This free space is, in part, occupied by students with a foreign background already applying to these schools. It should therefore be noted that these results on segregation do not necessarily generalize to other settings. In a case where the top-performing schools have a low share of students with a foreign background the effects could be smaller or, depending on the pre-existing choice pat-

Figure 1.10: Change in the share of students with foreign background in the oversubscribed schools loosing applicants



Note: Panel A shows the subscription rate of the middle schools in Linköping sorted by average performance. Panel B shows the share of the students with a foreign background in the two oversubscribed school that are not top-performing (1-5).

terns and the institutional setting, potentially even reversed.

To examine the extent to which these results are generalizable we use school data from the Swedish National Agency of Education. For all schools in all municipalities with ten or more middle schools (32 municipalities in total) we have information on each school's adjusted performance, share of foreign-born students and share of students with a foreign background. Within each municipality we assign each school a percentile rank based on its adjusted performance.[40] In Figure 1.12 in Appendix C we show binned scatterplots on how percentile adjusted performance ranks within these municipalities, covaries with the share of foreign-born students and students with a foreign background. The pattern we observed in Linköping, with the top-performing schools having a larger share of students with foreign background compared to the second-tier schools seems to be a general pattern. This implies that the results in this section should generalize, at least to other Swedish municipalities.

1.6 Concluding Remarks

In this paper we report results from a randomized controlled trial where information on absolute and adjusted (for student composition) school performance was distributed to households about to choose a middle school in a Swedish municipality. We find that this information affects school choice behavior. The effects are heterogeneous and mainly driven by high-skilled households with a native background. The changes in the choice behavior of these households also only seems to take place on the margin between the top-performing schools and the second-tier schools. Further, it seems that households are mainly reacting to the measure of adjusted per-

[40] Adjusted performance is defined as in our letter but, due to data limitations, using grades instead of test scores.

formance. This indicates that these households are not reacting to the information because they are looking for schools with a stronger (observable) peer composition. All in all this indicates that school performance information can have an effect on the competitive pressure on schools to improve performance. However, given that the effects are restricted to certain groups and certain schools, everyone might not benefit from this increased competition as disadvantaged students in low performing schools risk being left behind.

Using simulations, we find that the documented changes in school choice behavior translate into effects on the assignment of the households that did alter their choice behavior. Capacity constraints do however seem to mute this effect. This suggests that it is important to couple information interventions with a readiness to, in different ways, expand capacity at the schools for which demand increases. At least as long as it possible without creating other problems, The shift in choice behavior for high-skilled and native households also seems to displace students with a foreign background from the top-performing schools. Therefore, the intervention is not without drawbacks, at least as long as we care about equality, as it increases the school performance gap between advantaged and disadvantaged students. Finally, we find that initial effects on assignment propagates through the system as capacity frees up, thereby also potentially benefiting students that did not react to the information. Through this mechanism school segregation in terms of migration background is decreased.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Woolridge, J. (2017). When should you adjust standard errors for clustering? url: https://economics.mit.edu/files/13927 (accessed 1 April 2018).

Abdulkadiroglu, A., Agarwal, N., and Pathak, P. A. (2017). The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match. *American Economic Review*, 107(12).

Abdulkadiroglu, A., Pathak, P. A., and Walters, C. R. (2018). Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics*, 10(1):175–206.

Abdulkadiroglu, A. and Sönmez, T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747.

Allen, R. and Vignoles, A. (2007). What should an index of school segregation measure? *Oxford Review of Education*, 33(5):643–668.

Bifulco, R. and Ladd, H. F. (2007). School Choice, Racial Segregation, and Test-Score Gaps: Evidence from North Carolina's Charter School Program. *Journal of Policy Analysis and Management*, 26(1):31–56.

Böhlmark, A., Holmlund, H., and Lindahl, M. (2016). Parental Choice, Neighbourhood Segregation or Cream Skimming? An Analysis of School Segregation after a Generalized Choice Reform. *Journal of Population Economics*, 29(4):1155–1190.

Böhlmark, A. and Lindahl, M. (2015). Independent schools and long-run educational outcomes - evidence from Sweden's large scale

voucher reform. *Economica*, 82(327):508–551. CESifo Group Munich.

Burgess, S., Greaves, E., Vignoles, A., and Wilson, D. (2015). What Parents Want: School Preferences and School Choice. *The Economic Journal*, 125(587):1262–1289.

Carrington, W. J. and Troske, K. R. (1997). On Measuring Segregation in Samples with Small Units. *Journal of Business & Economic Statistics*, 15(4):402–409.

Corcoran, S. P., Jennings, J. L., Cohodes, S. R., and Sattin-Bajaj, C. (2018). Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions. *NBER Working Paper No 24471*. National Bureau of Economic Research, Inc.

Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The Effect of School Choice on Participants: Evidence from Randomized Lotteries. *Econometrica*, 74(5):1191?1230.

Duncan, O. D. and Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, 20(2):210–217.

Figalo, D. N. and Lucas, M. E. (2004). What's in a Grade? School Report Cards and the Housing Market. *American Economic Review*, 94(3):591–604.

Fiva, J. H. and Kirkeboen, L. J. (2011). Information Shocks and the Dynamics of the Housing Market. *The Scandinavian Journal of Economics*, 113(3):525–552.

Gallego, F. A., Lagos, F., and Stekel, Y. (2012). Types of information and school choice: An experimental study in the chilean voucher

system. url: http://www.webmeets.com/files/papers/LACEA-LAMES/2012/735/preliminary%20120312.pdf (accessed 1 April 2018).

Hart, C. M. D. and Figalo, D. N. (2015). School Accountability and School Choice: Effect on Student Selection Across Schools . *National Tax Journal*, 68(3S):875–900.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous preferences and the efficacy of public school choice. url: http://people.virginia.edu/ sns5r/microwkshp/hastings.pdf (accessed 1 April 2018).

Hastings, J. S. and Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *Quarterly Journal of Economics*, 123(4):1373–1414.

Hoxby, C. (2000). Does Competition Among Public Schools Benefit Students and Taxpayers? . *The American Economic Review*, 90(5):1209–1238.

Hsieh, C.-T. and Urquiola, M. (2006). The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's school Voucher Program. *Journal of Public Economics*, 90(8-9):1477–1503.

Hussain, I. (2013). Heterogeneous preferences and the efficacy of public school choice. url: http://www.sole-jole.org/13363.pdf (accessed 1 April 2018).

Imberman, S. A. and Lovenheim, M. F. (2016). Does the Market Value Value-added? Evidence from Housing Prices after a Public Release of School and Teacher Value-added . *Journal of Urban Economics*, 91(C):104–121.

James, D. R. and Taeuber, K. E. (1985). Measures of Segregation. *Sociological Methodology*, 15:1–32.

Kane, T. J. and Staiger, D. O. (2002). The Promise and Pitfalls of using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, 16(4):91–114.

Koning, P. and van der Wiel, K. (2013). Ranking the Schools: How School-Quality Information Affects School Choice in the Netherlands. *Journal of the European Economic Association*, 11(2):466–493.

Lavy, V. (2010). Effects of Free Choice Among Public Schools. *The Review of Economic Studies*, 77(3):1164–1191.

Lavy, V. (2015). The Long-Term Consequences of Free School Choice. *NBER Working Paper No. 20843*. National Bureau of Economic Research, Inc.

Lépine, A. (2015). School Reputation and School Choice in Brazil: a Regression Discontinuity Design. *Working Papers, Department of Economics 2015-38, University of São Paulo (FEA-USP).*

Lieberson, S. (1981). *A Piece of the Pie: Blacks and White Immigrants since 1880*. University of California Press, Oakland, California.

Massey, D. S. and Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, 67(2):281–315.

Mizala, A. and Urquiola, M. (2013). School Markets: The Impact of Information Approximating Schools' Effectiveness. *Journal of Development Economics*, 103(C):313–335.

Musset, P. (2012). School Choice and Equity: Current Policies in OECD Countries and a Literature Review. *OECD Education Working Papers, No. 66.* OECD Publishing, Paris, url: http://dx.doi.org/10.1787/5k9fq23507vc-en (accessed 1 April 2018).

Reardon, S. F. and Firebaugh, G. (2002). Measures of MultiGroup Segregation. *Sociological Methodology*, 31(1):33–67.

Ruijs, N. M. and Oosterbeek, H. (2017). School Choice in Amsterdam: Which Schools are Chosen when School Choice is Free? *Education Finance and Policy*, 0(ja):1–64.

Söderström, M. and Uusitalo, R. (2010). School Choice and Segregation: Evidence from an Admission Reform. *The Scandinavian Journal of Economics*, 112(1):55–76.

Theil, H. and Finizza, A. J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *The Journal of Mathematical Sociology*, 1(2):187–193.

Walters, C. R. (2014). The Demand for Effective Charter Schools. *NBER Working Paper No. 20640.* National Bureau of Economic Research, Inc.

White, M. J. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index*, 52(2):198–221.

Young, A. (2017). Channelling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. url: http://personal.lse.ac.uk/YoungA/ChannellingFisher.pdf (accessed 1 April 2018).

Appendix A - Complete Letter

Stockholms universitet

Further Information Regarding the School Choice

You have recently, or will soon, receive a letter from Linköping municipality informing you that it is time to choose a school for your child. As we are studying school choice, we would like to give you some additional information that we hope can be of use.

In the attached table you can see how the schools in Linköping municipality have performed on the standardised tests in grade nine in recent years. The first column shows the average score on the tests. In the second column we have calculated how the students of the school have performed on these standardised tests compared to schools that are similar in terms of the students' educational and migration background.

If you have any questions feel free to contact us.

Dany Kessel

dany.kessel@ne.su.se

0707173803

Elisabet Olme

elisabet.olme@ne.su.se

0735442220

Stockholms universitet

Skola School	Genomsnittspoäng Average Score (max=20)	Över/underprestation Over/underperformance (poäng/points)
Arenaskolan	-	-
Berzeliusskolan	14,4	+0,2
Björkö Friskola	13,2	-1,2
Dar al Uloum	-	-
Ekholmsskolan	12,6	-1,0
Elsa Brändströms Skola	-	-
Folkungaskolan	15,4	0
Frösunda Pandionskolan	-	-
Internationella Engelska Skolan	15,7	+0,6
Klara Privata Grundskola	13,5	+0,2
Kungsbergsskolan	11,7	-1,4
Kunskapsskolan Linköping	11,7	-2,0
Linghemsskolan	12,8	-0,8
Ljungsbro Skola	12,5	-0,8
Malmslättskolan Tokarp	13,7	-0,3
Nya Munken	14,2	-0,5
Skäggetorpsskolan	9,8	-0,8
Tornhagsskolan	13,1	-0,1
Vittra Lambohov	16,1	+1,0
Ånestadsskolan	13,0	+0,1

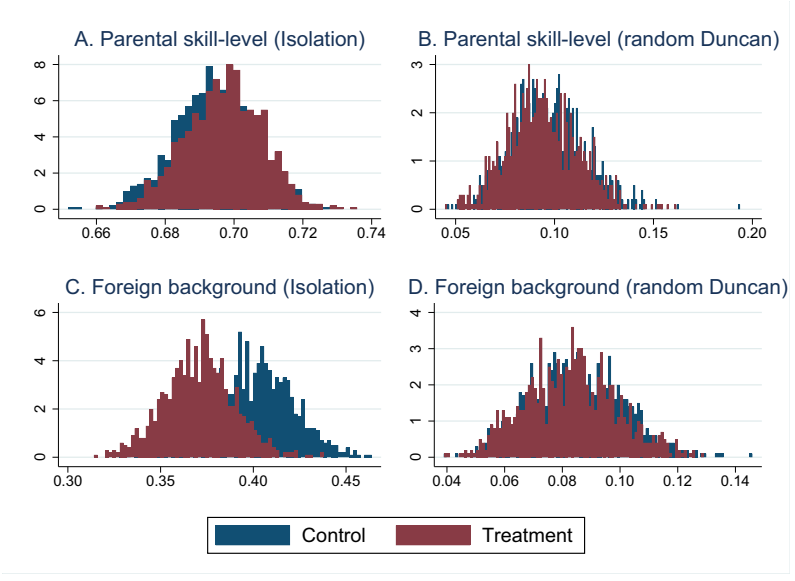
Såhär läser du tabellen: Om det står 15 i kolumnen ”Genomsnittspoäng” och +0,2 i kolumnen ”Över/underprestation” betyder det att genomsnittspoängen för skolans elever var 15 och att detta var 0,2 poäng mer än genomsnittet på skolor med liknande elever.

How to read the table: If it says 15 in the column “Average performance” and +0,2 in the column “Over/underperformance”, it means that the students average score at the school was 15, and that this was 0,2 points higher than the average score at schools with similar students.

Källa/Source: Skolverket/The National Agency for Education.

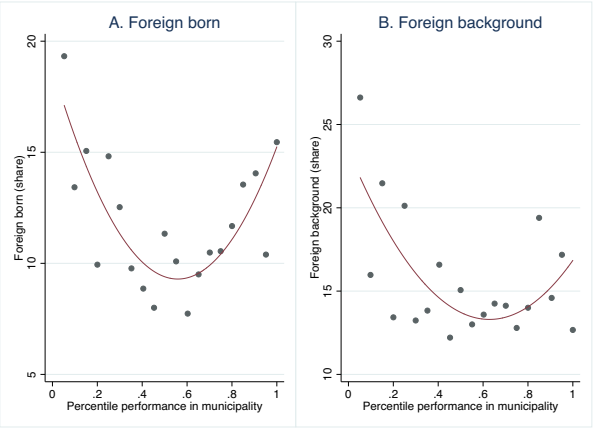
Appendix B - Segregation, Robustness

Figure 1.11: Effects on segregation, robustness



Appendix C - Performance and Background in Sweden

Figure 1.12: Adjusted performance and foreign background in Sweden



Note: This figure shows the relationship between a schools relative adjusted performance (within municipality) and the share of foreign-born students and students with a foreign background.

Chapter 2

School Choice Priority Structures and School Segregation*

*This chapter was co-authored by Elisabet Olme. We would like to thank Chris Nielsen, David Strömberg, Jonas Vlachos and seminar participants at Stockholm University, Stockholm-Uppsala Education Economics Workshop, Columbia Business School, Columbia University, Princeton University, INAS conference Oslo, CEN workshop Copenhagen, SUDSWEC, IAS Norrköping, Lund University, Stockholm School of Economics, AASLE, Reus Education workshop, Tilburg University and Toulouse ENTER conference for valuable comments and feedback. We also gratefully acknowledge financial support from *Ge alla elever samma chans* (a project aimed at reducing the inequality in the Swedish educational system run and funded by The Swedish Trade Union Confederation (LO) and the two major Swedish teachers unions (Lärarnas Riksförbund and Lärarförbundet)).

2.1 Introduction

Traditionally, in many school systems, children about to start primary school have simply been assigned to their neighborhood schools. In recent decades, school choice programs have rapidly been introduced all over the world. This has lead to a sharp increase in the number of households that, each year, are faced with the decision of where to send their children to school.[2] Empirical evaluations seem to indicate that introducing or expanding school choice opportunities leads to increased school segregation in various dimensions.[3] It is clear that school segregation is a big concern in the public debate, in Sweden and elsewhere. School segregation has also been shown to have adverse effects on student outcomes, labor market outcomes and crime.[4] A central question is therefore why school choice tends to increase segregation? The introduction of school choice enables households to apply to their preferred schools. As preferences for school attributes seem to vary systematically with observable household characteristics, one explanation could be that school choice allows households to self-segregate according to these preferences.[5]

The outcome of a school choice program is however not only a function of the households' preferences, but also of the design of the program. School choice programs can take many different forms, they can be general programs including all students and schools within a

[2]One could say that there is a school choice component in a household's decision on where to reside. However, when referring to the term school choice in this paper we mean a system in which households can express preferences for schools and these preferences are taken into consideration in the assignment process.

[3]See Hsieh and Urquiola (2006), Bifulco and Ladd (2007), Söderström and Uusitalo (2010) and Böhlmark et al. (2016).

[4]See e.g. Hanushek et al. (2000), Billings et al. (2013), Johson (2015), Gamoran and An (2016) and Billings et al. (2016).

[5]See e.g. Hastings et al. (2009), Borghans et al. (2015), Burgess et al. (2015) for papers documenting heterogeneity in preferences for school attributes.

specific geographic area, they can target specific groups of students or include only specific types of schools. No matter the scale of the program, a necessary component is the priority structure - a rule determining admission to oversubscribed schools. We ask if the seemingly segregating effects can be mitigated by altering these priority structures and whether different types of priority structures benefit different types of households? To study this, we turn to the mechanism design literature.

The first paper to address the school choice problem from a mechanism design perspective was Abdulkadirolu and Sönmez (2003). They concluded that there are flaws in many commonly used school assignment mechanisms and proposed two alternative mechanisms - The Deferred Acceptance (DA) and The Top Trading Cycles (TTC) mechanism. Since then, there has been a growing academic interest in the design of school choice programs. The main focus has been on the assignment mechanism as such and the properties of the allocations they produce. This discussion has led to the adoption of new assignment mechanisms, most famously in Boston and New York.[6] Less attention has been given to the importance of the priority structure. As pointed out by, among others, Fack et al. (2017), *"Despite the fact that admission criteria impact student sorting and student welfare ... they have been relatively under-studied in the literature"*.[7] We study how different priority structures impact school segregation and student welfare in a school choice program using a modified version of the DA algorithm to assign students to public schools.[8]

[6]Abdulkadirolu et al. (2005b) discuss the assignment mechanism used in Boston until 2005 and their recommended alternatives and Abdulkadirolu et al. (2009) discuss the redesign of the assignment process for NYC high schools in 2003-2004.

[7]There are some exceptions, for example, Calsamiglia and Güell (2014) find that when proximity-based priorities are used in the Boston mechanism, the student allocation is very similar to a neighborhood-based allocation.

[8]The DA algorithm is becoming increasingly popular and is now used in e.g.

We observe all school choices of primary school starters in the Swedish municipality of Botkyrka from 2011 to 2014. The focus on primary school starters is motivated by the fact that many children in Botkyrka attend the same school for their entire elementary education, making the choice of primary school an important one. The Botkyrka School Choice Program uses a modified version of the DA mechanism combined with a priority structure based on proximity and already enrolled siblings to assign students to public schools. The school choice process is centralized and includes all primary schools in the municipality.[9] In order to perform an ex ante policy evaluation, we rely on a structural approach.[10] Using a rank ordered mixed logit model and observed school choices we estimate parental preferences for school attributes. These preferences are used to construct school rankings for artificial cohorts of students. These students and their imputed school rankings are used in simulations to evaluate the impact of different priority structures. We record the level of segregation, and a number of proxies for student welfare (such as the average distance from home to the assigned school and the share of students that are admitted to their most preferred school) for the whole population as well as for different subgroups. Our empirical strategy is similar to the one used by Gallego and Hernando (2008) and Fack et al. (2017).

Our simulations rely on school rankings generated using estimated preferences for school attributes. In order for these rankings to be valid in our counterfactual scenarios we need households to be truth-telling when submitting their school choices. In the general case,

Amsterdam, Boston, Paris and New York (Fack et al., 2017) and is one of the assignment mechanisms recommended by Abdulkadirolu and Sönmez (2003).

[9]Voucher schools are excepted from using the DA mechanism and the municipal priority structure.

[10]Pathak and Shi (2017) suggest that the use of structural demand models in the school choice context can be effective when studying counterfactual outcomes.

the deferred acceptance mechanism is strategy proof. Haeringer and Klijn (2009) do however show that putting a restriction on households regarding how many schools they can rank, reintroduces strategic incentives. In Botkyrka, as in many other places, there is such a restriction (three schools in Botkyrkas case). We evaluate the size of this problem by using the concept of "safe schools", i.e. schools where a household can be sure to get admitted if they apply to it.[11] We show that if a household has ranked a "safe school" as their top or second choice, they can be assumed to be truth-revealing. Due to the significant overcapacity that Botkyrka's school choice program is operating under during this time period we can categorize 95 percent of households as truth-revealing.

We evaluate three different priority structures. The first priority structure we impose gives priority to students according to the distance between the students' homes and the school in question (proximity). The second priority structure assigns each student a randomly drawn lottery number and gives priority based on that. The last priority structure is designed to contribute to the socioeconomic and ethnic diversity at the schools, and falls under what Abdulkadirolu and Sönmez (2003) call controlled choice. At each school, s_j seats are reserved for students of type j, where s_j equals the share of type j students in the current cohort of primary school starters, times the total capacity of the school. Within type, priorities are determined by distance between the students' homes and the school. Similar types of priority structures are currently in use in actual school choice programs.

As long as proximity is not the only school attribute that parents care about, we would expect that a system where seats are reserved for underrepresented groups will produce a lower level of school segregation compared to a system where priority is based on proximity.

[11] A concept organically proposed in Calsamiglia et al. (2010).

How a system using lottery-based priorities will perform in terms of school segregation is, ex ante, less clear. It will remove the privilege of those residing near popular schools, it does however also make it easier for students to opt out of nearby schools they don't like. Hence, depending on the search patterns and the level of residential segregation, this could both increase or decrease the level of segregation than a system where priorities are based on proximity. In terms of welfare, the average distance to the assigned school is likely to increase whenever priorities are not based on proximity. Given the importance most households seem to put on proximity, this is a potential welfare loss. However, an increase in distance could also reflect the fact that households that are willing to travel further to get into a more preferred school, can now do this. In which case the increased home to school distance would reflect an increase, rather than a decrease in welfare. When it comes to the share of students assigned to their most preferred schools, we have less clear predictions.

The paper most similar to ours is Fack et al. (2017) that studies the effects on student sorting when altering the priority structures used in the Paris high school choice program. In their empirical analysis they focus mainly on the effects on ability sorting and the welfare effects on high- and low-ability students. They find that there are clear trade-offs between sorting and aggregate welfare and between the welfare of high and low-ability students. They briefly look at sorting in terms of socioeconomic background and find that a priority structure where low-income students are given a "bonus" in terms of priority, produces a lower level of sorting compared to systems based on lotteries or grades. We contribute to these findings in three ways. Firstly, we examine the question in a different setting, specifically one where children are younger. This is relevant as mobility varies with age, with older children being more likely to consider traveling further to school. Hence, the priority structure could be

less important when children are younger. Secondly, and somewhat related to the first point, we examine other priority structures. Fack et al. (2017) use the default system in Paris, where students who receive means-tested low-income financial aid (15 percent of students) get priority to their preferred school. We use a system specifically designed to minimize segregation, both in terms of parental education and migration background. Thirdly, we have access to high-quality Swedish register data and can therefore perform detailed subgroup analysis.

The results suggest that using reserved seats for children from different socioeconomic groups produces a lower level of school segregation than proximity- or lottery-based priorities. Segregation in terms of foreign background is about four points lower as measured by the Duncan Dissimilarity Index (two points measured by Theil Entropy Index) and segregation in terms of parental eduction is two/one point(s) lower. When compared to a system based on proximity, this reduction in segregation comes with a small increase in the average distance from the students' homes to their assigned school (40 meters from a baseline of 957 meters) and a minor decrease (less than one 1 percentage point from a baseline of 83 percent) in the share of students being admitted to their most preferred school. The increase in average distance to school is however driven by students getting their most preferred school and it is therefore unclear if this should be considered as an actual cost. The lottery-based system is underperforming compared to the two other systems, producing both a larger average distance to school and a lower share of students admitted to their most preferred school. Even though the system based on proximity and the one based on reserved seats, on average, produce similar results in terms of distance to school and the share of students getting admitted to their top choice, there are clear winners and losers of the two systems. Highly-educated households

with a native background benefit from the system based on proximity, while low educated households with a native background and highly-educated households with a foreign background, benefit from a system based on reserved seats.

The paper is outlined as follows. Section 2.2 presents the institutional setting followed by Section 2.3 describing the data and presenting summary statistics. Next, Section 2.4 discusses the priority structures that will be used to study counterfactual outcomes and Section 2.5 outlines the empirical strategy. In Section 2.6 we present our results. Section 2.7 concludes.

2.2 Institutional Setting

The Swedish elementary school system covers ten grades with children usually starting school the year they turn six. The first year (grade K) is voluntary while the remaining nine years are mandatory (grade 1-9). Despite the first year being voluntary, almost everyone (97 percent) attends. Since the early 1990s municipalities are responsible for providing and financing the primary education of all children residing in the municipality. There are two types of schools: public schools that are run by the municipalities and voucher schools that are operated by independent providers, but publicly funded. All schools follow the same curriculum, stipulated in national legislation. School funding is based on vouchers directly connected to each student but municipalities are allowed to adjust the amount depending on a student's background or disabilities.[12] Neither public nor voucher schools are allowed to charge any fees.

School choice was formally introduced in 1992, along with the reform that allowed voucher schools to operate. Prior to this, students,

[12]In Botkyrka, 28-44 percent of the school budget is compensatory (Ernst & Young AB, 2014).

in general, attended their local public school. Initially, school choice was decentralized to the school level requiring parents that did not want their children to attend their local public school to contact other schools directly and see if there was a vacancy for their child. Naturally, the vast majority continued to attend their local public school. At the turn of the millennium the expansion of voucher schools took off.[13] Partly in reaction to this expansion, municipalities, beginning in the mid-2000 started to implement centralized school choice programs. Most municipalities, at least in the urban areas, now have centralized school choice programs where households can apply to all public schools in their municipality. In addition, households are allowed to apply to all voucher schools, both within and outside their own municipality. It is also possible to apply to public schools in other municipalities, but these schools have no obligation to accept students not residing in their municipality.[14]

As primary education in Sweden is decentralized, there is considerable variation in the design of the school choice programs. There is some national legislation regulating how priorities to oversubscribed schools should be determined. Municipalities have to allocate students to schools based on their expressed preferences while at the same time making sure that no students are admitted to a school too far away from their homes.[15] Voucher schools are restricted to accepting students based on a first-come first-served basis, already enrolled siblings and the distance from the students home to the school. In practice, voucher schools seldom use the last ground for

[13]Böhlmark and Lindahl (2015) show that the share of the students enrolled in voucher schools increased from 1.6 percent in 1998 to 11 percent by 2009.

[14]Under special circumstances, a child can have the right to attend a public school located in another municipality (`http://www.skolverket.se/regelverk/mer-om-skolans-ansvar/val-av-grundskola-1.210176`).

[15]Children always have the right to stay in a school they are already enrolled in, at least as long as it is located in the municipality they live in.

admission but give priority based on enrolled siblings and time in the school-specific queue.[16]

2.2.1 The Botkyrka School Choice Program

Botkyrka is the fifth largest municipality in Stockholm county. It has a diverse population of 90,000 inhabitants, 56 percent of which have a foreign background.[17] There are 23 primary schools offering grade K in Botkyrka, about two thirds of these cover all years up to grade 9. Given this, most children stay in the same school for their entire elementary education. Although most schools are public, there are five voucher schools that together account for about 8 percent of the students.[18] Figure 2.1 displays the location of all schools that were available to choose from in 2014, as well as the number of primary school starters residing in each SAMS[19] area in the municipality. The schools are concentrated to three different neighborhoods, one in the north, one in the east and one in the west. Descriptive statistics of both schools and primary school starters will be presented for the whole population as well as for each neighborhood.

The average cohort of primary school starters in Botkyrka during this time period consists of just under 1,000 students. About 95 percent of households with a child about to start primary school participate in the municipality's centralized school choice program.

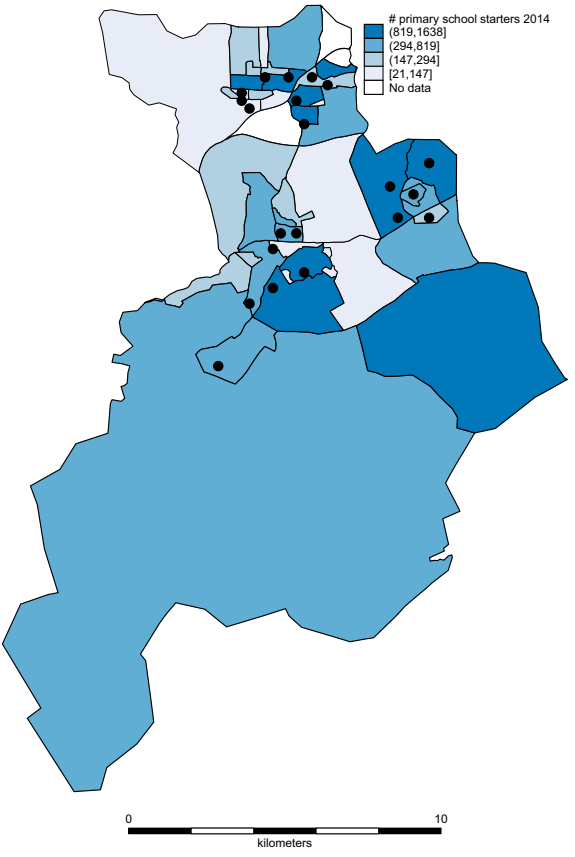
[16]The regulations can be found in Swedish at `http://www.skolverket.se/regelverk/mer-om-skolans-ansvar/val-av-grundskola-1.210176`.

[17]A person is categorized as having a foreign background if he/she is born abroad and/or if both his/her parents are born abroad.

[18]Four of the voucher schools are likely to cater to a very specific group of households. Two are religious (one Christian and one Muslim), one is bilingual (Finnish-Swedish) and one uses a specific teaching method inspired by Freinet.

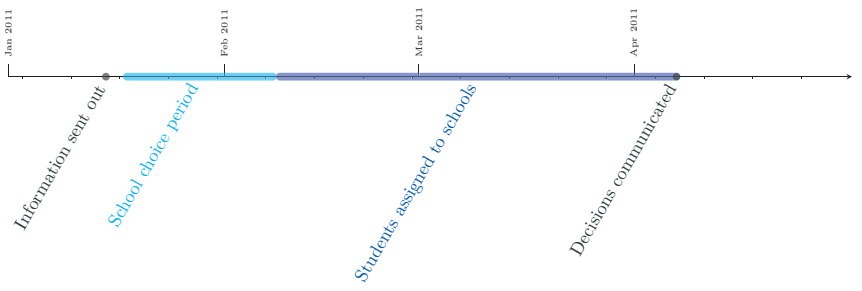
[19]SAMS stands for Small Areas for Market Statistics and is a partitioning of Sweden into about 9,500 smaller areas, based on municipal partitioning and electoral districts, depending on the size of the municipality.

Figure 2.1: Map of school location and number of primary school starters in Botkyrka



Note: This map is generated using all primary school starters in 2014 and the primary schools available for them to choose from. Each SAMS area (©SCB) in the municipality is color coded by the number of primary school starters residing in that area.

Figure 2.2: Timeline of the Botkyrka School Choice Program



The timeline of this program is illustrated in Figure 2.2. In January, all households with a child expected to start grade K in the upcoming academic year are sent a letter informing them about the school choice program. The letter also contains information about both the public and voucher schools in the municipality. Then there is a three week period during which parents can register their school choices using an online tool. All parents have to submit three ranked choices from a list of all public and voucher schools in the municipality. Households can also state that they have applied to a school outside the municipality, although they cannot send their application to these schools through the online tool. About three percent of the households use this option. Once the school choice period ends, students are allocated schools and in April households are informed about which school they have been assigned to. Participation is encouraged by there being no default school where a child is guaranteed a seat if the household does not participate in the program. This gives strong incentives to submit an application list.

When allocating student to public schools the municipality used a modified version of the Deferred Acceptance algorithm. Voucher schools are excepted from this and prioritize students that list them as

their top choice, on a first-come first-served basis and already enrolled siblings. In practice, this implies that admission takes place in two rounds, beginning with the admission to the voucher schools followed by admission to public schools (Where students already admitted to voucher schools are excluded). This is however not something households are informed about and in data we observe households ranking voucher schools as their second or third most preferred school as well.

The deferred acceptance algorithm was proposed by Gale and Shapley (1962) and is becoming an increasingly more common mechanism to assign students to schools. Different versions of this algorithm are used in Boston, New York, Amsterdam and Paris (Fack et al., 2017). The following description of the algorithm is based on Abdulkadirolu and Sönmez (2003). Assume that all students have strict preferences over all schools and that all schools have a strict priority ordering of all students.

Step 1 Each student applies to his/her top choice. Each school tentatively accepts one student at a time according to their priority ordering until their capacity is reached or there are no more applying students.

Step k Each student rejected in round k−1 applies to the next school in their ranking. Each school considers all tentatively accepted and new applicants together and tentatively accepts one student at a time, according to their priority ordering, until their capacity is reached or no applicants remain.

The algorithm terminates when no student is rejected, at which point all students are placed at their final assignment. In the deferred acceptance algorithm, a student who has been tentatively accepted

at a school can be rejected from the same school in a later round, as priorities are independent of how a student ranks the school. This property of the DA algorithm ensures that it is strategy proof, as shown by Abdulkadirolu and Sönmez (2003).[20] Botkyrka deviates somewhat from this procedure, as households are only allowed to rank three schools. Haeringer and Klijn (2009) show that this reintroduces strategic incentives. When rankings are truncated, households may want to "play it safe" and include less preferred schools in order to avoid a situation where the student is rejected from all the schools they have applied to. Another important point, made by Fack et al. (2017), is that the strategy proofness of the DA mechanism only implies that being truthful is a *weakly* dominant strategy. We will discuss the concept of truthfulness and its consequences for this study in more detail in Section 2.5.3.

The priority orderings of public schools in Botkyrka are based on proximity and already enrolled siblings. Children that have a sibling enrolled in grade K-4 at a school have priority to that school over children that do not. To get a strict priority ordering, children with and without siblings are ranked using a relative distance measure meant to minimize the walking distance of all children. It is defined as the distance to the nearest school (except the one for which one is calculating the relative distance to) minus the distance to the school in question. A higher relative distance measure gives higher priority.[21] Children in households that do not submit a ranked list of schools

[20]The deferred acceptance is also stable (eliminates justified envy), which is not the case for the other strategy proof algorithm proposed by Abdulkadirolu and Sönmez (2003) (TTC). On the other hand, the TTC is Pareto-efficient which is not the case for DA. In practice, school officials face a trade-off between stability and efficiency and the increasing popularity of the DA suggest that stability is given more weight than efficiency in these situations.

[21]The distance is the walking distance measured using only roads that the municipality have assessed to be safe for a 6-year old to walk along. This is to ensure that children do not need to be accompanied by an adult to get to school.

and children that are not admitted to any of their chosen schools are assigned to schools where there is free capacity once all other children are assigned to schools.

2.3 Data

From Botkyrka municipality we have access to all school choices registered in the online tool from 2011 to 2014. In total we observe the school choices of 3,820 primary school starters. The data includes the three ranked choices of each household. Using individual identifiers we can link these choices to register data provided by Statistics Sweden. Further, using the multigenerational register, we can link the primary school starters to their parents and siblings. In addition, we have access to data on all students enrolled in a primary school in Botkyrka during this time period, allowing us to create measures of the student composition at each school.[22] This is complemented with data on school performance, survey data on what students and parents think about their schools and teacher characteristics. Tables 2.6 and 2.7 in Appendix A provide definitions of all variables at the individual and school level.

Table 2.1 reports summary statistics of the pooled population of primary school starters 2011-2014. We can see that there are large differences between neighborhoods. The North is characterized by a high fraction of students with foreign background and low-educated parents, whereas the opposite is true for the East. In all areas, about half of the primary school starters have an older sibling enrolled in a primary school in Botkyrka. Once the primary school starters have

[22]We do not observe the students attending the two confessional schools, Statistics Sweden do not provide information on students in religious schools in order to protect their religious rights. Consequently, even though there are 23 schools available to choose from in Botkyrka our analysis will only be based on 21 of these.

enrolled, they have on average a distance of about 1,000 meters to school. Primary school starters in the West travel a bit further than those in the two other neighborhoods, which is probably due to the fact that this neighborhood includes the more rural parts in the southwest of Botkyrka.

In Table 2.2, summary statistics of the available schools (at the school-year level) are reported. First, we note that the differences in characteristics between the primary school starters in the three different neighborhoods are translated into differences in the schools' student compositions. Schools in the North have a much higher fraction of students with a foreign background and less educated parents. The average score on the standardized tests is 13.8 points on a scale from 0 to 20, with the North performing below average and the other two neighborhoods performing above average. In all neighborhoods, the average capacity in grade K is about the same (55 seats). Additional variables on the school level that are used in robustness tests are presented in Table 2.9.

Table 2.3 shows that, on average, households have about five schools within two kilometers from where the live. Increasing the distance from two to three kilometers gives two additional schools to choose from. As we know, from previous research, that proximity of the school is an important factor in school choice, this is important since it establishes that there is a real choice of schools even for households that have strong preferences for proximity. Table 2.3 also suggests that there are significant differences in school attributes even for households restricting their choice set to schools close to their home. For example, for the average households, the school within 2 kilometers of their home, with the highest fraction of students with a foreign background, has a 22 percentage points larger share of students with a foreign background compared to the school within 2 kilometers of their home with the lowest share of

students with a foreign background. Looking at test scores we can see that the average household can improve by almost two points by choosing the highest performing instead of lowest performing school within two kilometers of their home. In other words, if households care about these attributes they have strong incentives to participate in the school choice program even if proximity restricts the set of schools relevant to consider.

In Table 2.4 we report some summary statistics of the school choice behavior in our sample, focusing on the location of the chosen schools. Even though proximity is likely to be an important determinant of school choice, it is evident that some households are willing to travel in order to get into a school that better matches their preferences. Seven percent of all households list at least one school outside their own neighborhood among their choices. The proportion of households trying to opt out of their neighborhood is higher in the North and West than the East. Note also that the fraction applying to a voucher school is much higher in the North than in the other two neighborhoods. This is most likely due to the fact that four of the five voucher schools are located in the North (one is located in the West). Finally, a very high fraction of all primary school starters are assigned their top choice. We observe that 90 percent of all students attend their top choice. Note that this is a lower bound as choices are registered in January while attendance is measured in October and during these months households might have changed schools (for example because they changed their minds or because they relocated). This high number is most likely due to the fact that the municipality has a significant overcapacity (10-20 percent) in its school system.[23]

[23]The overcapacity observed in Botkyrka during the time period studied is explained by the fact that these four cohorts of primary school starters were relatively small compared to the cohorts before them. The municipality was expecting future cohorts to increase in size and therefore thought a shorter period of overcapacity in the schooling system was motivated.

Table 2.1: Summary statistics of primary school starters by neighborhood

	Total mean/sd	North mean/sd	West mean/sd	East mean/sd
Native, highly-educated	0.34	0.09	0.38	0.67
	(0.47)	(0.29)	(0.49)	(0.47)
Native, low-educated	0.25	0.17	0.35	0.24
	(0.43)	(0.38)	(0.48)	(0.43)
Foreign, highly-educated	0.18	0.29	0.15	0.06
	(0.39)	(0.45)	(0.36)	(0.24)
Foreign, low-educated	0.23	0.45	0.11	0.03
	(0.42)	(0.50)	(0.31)	(0.18)
Sibling in the school system	0.49	0.48	0.51	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
Distance to attended school	1027.93	880.41	1362.41	788.76
	(1313.60)	(1217.64)	(1646.54)	(694.88)
Distance to top choice	1051.14	925.04	1383.17	783.01
	(1332.84)	(1268.26)	(1647.61)	(696.30)
No of schools in nbd	7.30	9.00	7.00	5.00
	(1.61)	(0.00)	(0.00)	(0.00)
Attending voucher school	0.08	0.18	0.03	0.01
	(0.27)	(0.38)	(0.16)	(0.07)
Observations	3820	1564	1260	996

Note: This table presents the mean and standard deviation of each variable. The sample includes all primary school starters choosing schools in Botkyrka 2011-2014. *Highly-educated* indicates that at least one parents has an education level above upper secondary school. *Foreign* indicates that the student was born abroad and/or that both parents born abroad, otherwise the student is categorized as *native*. *Sibling in the school system* indicates if there is a sibling currently attending grade K-9 in any primary school in Botkyrka. *Distance to attended school* and *top choice* measures the home to school distance to the attended school and the school listed as the top choice in meters. *No of schools in the neighborhood* counts how many schools that are available to choose from within the student's own neighborhood. *Attending voucher school* is the share of students enrolling in a voucher school.

Table 2.2: Summary statistics of schools by neighborhood

	Total mean/sd	North mean/sd	West mean/sd	East mean/sd
Average test score	13.84	13.27	14.03	14.59
	(1.00)	(0.94)	(0.91)	(0.57)
Foreign background	0.48	0.80	0.32	0.12
	(0.31)	(0.09)	(0.18)	(0.03)
Highly-educated parents	0.46	0.32	0.49	0.67
	(0.16)	(0.07)	(0.10)	(0.09)
School capacity	54.94	55.00	55.09	54.60
	(18.15)	(12.62)	(24.15)	(17.97)
Observations	84	36	28	20

Note: This table presents the mean and standard deviation of each variable. The sample includes one observation per year of all schools available for choice in Botkyrka 2011-2014. *Average test score* is the school's average score (max = 20) on the standardized tests in Swedish and mathematics taken in grade 3. *Foreign background* indicates the share of students born abroad or with both parents born abroad. *Highly-educated parents* indicates that at least one parent has attended higher education. *School capacity* is the average number of seats available at each school for primary school starters.

Table 2.3: Number of schools and difference in attributes by home to school distance

	0.5km mean/sd	1km mean/sd	2km mean/sd	3km mean/sd	5km mean/sd
Number of schools	0.54	1.91	4.93	6.56	11.61
	(0.60)	(1.25)	(2.15)	(2.24)	(3.71)
Difference (max-min)					
Share foreign background	0.01	0.11	0.22	0.28	0.68
	(0.04)	(0.11)	(0.14)	(0.16)	(0.18)
Share highly-educated	0.00	0.09	0.19	0.24	0.44
	(0.02)	(0.08)	(0.08)	(0.08)	(0.11)
Test scores	0.09	0.88	1.72	1.89	2.79
	(0.34)	(0.84)	(0.88)	(0.86)	(0.82)
Observations	3820	3820	3820	3820	3820

Note: This table displays the mean and standard deviation of each variable. The sample includes all primary school starters during 2011-2014. *Number of schools* indicates how many schools the average student has to choose from within x kilometers. The remaining variables display the difference between the maximum and minimum value of the schools within x kilometers in terms of *Foreign background*, *Highly-educated parents* and *Test scores*.

Table 2.4: School choice behavior by neigborhood

	Total mean/sd	North mean/sd	West mean/sd	East mean/sd
Outside neighborhood				
Top choice	0.04	0.06	0.03	0.01
	(0.18)	(0.23)	(0.18)	(0.09)
Any choice	0.07	0.09	0.08	0.02
	(0.25)	(0.29)	(0.27)	(0.13)
Two or more choices	0.04	0.06	0.03	0.01
	(0.18)	(0.23)	(0.17)	(0.08)
Outside Botkyrka				
Any choice	0.03	0.05	0.01	0.00
	(0.16)	(0.22)	(0.12)	(0.04)
Voucher				
Any choice	0.15	0.31	0.06	0.01
	(0.35)	(0.46)	(0.23)	(0.08)
Attends				
Top choice	0.90	0.85	0.94	0.94
	(0.29)	(0.36)	(0.24)	(0.23)
No choice	0.05	0.09	0.02	0.01
	(0.21)	(0.28)	(0.13)	(0.12)
Outside neighborhood	0.03	0.05	0.03	0.01
	(0.17)	(0.21)	(0.17)	(0.08)
Observations	3820	1564	1260	996

Note: This table displays the mean and st. deviation of each variable. The population includes all primary school starters during 2011-2014. The first three variables indicate whether *top choice*, *one choice* or *two or more choices* are schools located outside one's neighborhood. *Any choice outside Botkyrka* is an indicator for listing at least one school located outside the municipality. *Voucher, any choice* is an indicator for listing at least one voucher school. *Attends top choice* is an indicator for attending the top choice school and *no choice* indicates that the student was not assigned to any of the three schools applied to. *Outside neighborhood* indicates that the student was assigned a school outside it's own neighborhood.

2.4 Counterfactual Scenarios

As mentioned, we evaluate three different priority structures. In this section we describe these three alternative priority structures.

Proximity-based priorities: The first priority structure we impose gives priority to students according to how close to the school they live. Each school ranks all students according to how far away from the school they live, giving priority to those residing closer to the school. The distance is calculated as the straight line between the student's residence and the school's location. Proximity-based priorities are common in Sweden, although recently municipalities have started to use different kinds of relative distance measures, rather than the absolute distance to a school. They are also used in the UK (Burgess et al., 2015)

Lottery-based priorities: This priority structure assigns each student a randomly drawn lottery number. Each school ranks all students according to this lottery number, giving priority to those with a lower number. The lottery number determines the student's priority at every school, i.e. the student with lottery number 1 will have the highest priority to all schools, the student with lottery number 2 will have the second highest priority at all schools and so on. This follows Abdulkadirolu et al. (2009), showing that this way of assigning lottery numbers is favorable in terms of student welfare compared to a scenario where each student is given a separate lottery for each school. Note also that with this specific priority structure, the DA mechanism is equivalent to the random serial dictatorship (also described in Abdulkadirolu and Sönmez (2003)). Lottery-based admission is used in e.g. Amsterdam, Beijing and New York city (see De Haan et al. (2015), He (2017) and Abdulkadirolu et al. (2005a)).

Proximity-based priorities with reserved seats: The last priority structure is designed to contribute to the socioeconomic and ethnic diversity at the schools, and falls under what Abdulkadirolu and Sönmez (2003) call controlled choice. In their version, it consists of (racial) quotas that are imposed on the schools. However, Kojima (2012) shows that this implementation of affirmative action in school choice can actually hurt rather than help the students of the type meant to benefit from the policy. Using simulations, Hafalir et al. (2013) confirms that these results are not an exception but rather the rule. Instead, they suggest a different implementation of affirmative action in school choice called minority reserves, where minority students are given higher priority as long as the seats reserved for the minority are not full. We follow their implementation of affirmative action and define four student types based on foreign background and parental education. At each school, s_j seats are reserved for students of type j, where s equals the share of type j students in the current cohort of primary school starters, times the total capacity of the school. Within type, priorities are determined by distance from the students home to school. In each round, if there are less applicants of type j than the number of seats reserved for this type, s_j, those seats will be open for students of other types. For a more detailed description of how assignment with reserved seats is performed, see Appendix D. There are several examples of school choice programs where affirmative action has been incorporated. One such example is the high school system in Paris, using quotas for certain groups in order to give disadvantaged students access to better high schools and reduce school segregation (Fack et al., 2017). In the US, such policies are sometimes the result of court ordered desegregation guidelines (Abdulkadirolu and Sönmez, 2003).

Note that a student's priority at a school is not dependent on how that student ranked the school. That is true for all three priority structures under study. Further, in our simulation we construct complete school rankings for all students. These facts combined with our use of the (student-proposing) DA mechanism guarantee that there are no strategic incentives in our counterfactual scenarios. Note also that the alternative priority structures are imposed on public as well as voucher schools.

2.5 Empirical Strategy

To evaluate the impact of alternative priority structures we rely on estimated preference parameters to create complete rankings of all schools for all students. This section describes the simulation strategy as well as the discrete choice framework used to estimate the preference parameters.

2.5.1 Simulations

We draw a random sample of 1,140 primary school starters, from our pooled population 2011 to 2014, retaining their observed background characteristics and residential location.[24] For each primary school starter we construct a complete ranking of all schools available to choose from, using estimated preference parameters for different school attributes (see section 2.5.2 for details) to calculate their utility according to the following specification of student i's utility from

[24]We choose 1,140 as the size of our artificial cohorts in order to limit overcapacity in the system. As mentioned, Botkyrka, under this time period operates with a large overcapacity in its school choice system. A large overcapacity makes the priority structures irrelevant. In the long run it is also not sustainable. Therefore, in our counterfactuals we choose a cohort size that better matches the total capacity (2 percent overcapacity).

attending school j

$$U_{ij} = \boldsymbol{X'_{ij}}\boldsymbol{\beta_i^x} + \epsilon_{ij} \tag{2.1}$$

where $\boldsymbol{X}_{ij}$ is a vector of school characteristics that vary at the school or school-individual level. The preferences for these attributes are captured by $\boldsymbol{\beta}_i^x$ and ϵ_{ij} is a random component that is assumed to be iid with a type 1 extreme value distribution. We allow households to have preferences for a number of school characteristics. These include: school performance (measured by the average pass rate on standardized tests in grade 3), the share of students with highly-educated parents, the share of students with foreign background, the distance between the school and the student's home, if the school is located in the same neighborhood as the household and whether the household already has other children attending that school. Further, we assume that households consider the latest available information on school attributes and expect them to stay the same. We think that this is a plausible assumption about how parents inform themselves and make their school choices, especially since the educational market in Botkyrka municipality is quite stable across the time period of study. Figure 2.11 in Appendix B shows the number of students, the share of students with highly-educated parents, the share of students with a foreign background and the average test score for each school and each year and suggests that school attributes are stable during the relevant time period.

For our first randomly drawn sample, each household's utility for each school is calculated using the school type characteristics as observed in the last year of our data, 2014. Based on these utility calculations we create complete rankings over all schools for each households. The primary school starters are allocated to schools using deferred acceptance and one of the priority structures under evaluation. Once allocated, the student composition of the schools is

updated according to the recently entered cohort and a new sample is randomly drawn and allowed to enter the system. We use 100 iterations and study the evolution over a ten year period. As in Fack et al. (2017), the simulations are dynamic in the sense that we allow each new cohort of households to rank the schools based on school attributes that are updated given the allocation of previous cohorts.

We are interested in a number of outcomes. The primary outcome of interest is the level of school segregation for the entering cohort. We study segregation along two dimensions, parental education and foreign background. Segregation between groups divided into organizational units (such as schools) can be measured using either measures of evenness or measures of exposure (Massey and Denton, 1988). Exposure measures are sensitive to the share of minority students in the population, which evenness measures are generally not. As Allen and Vignoles (2007) point out the share of minority students is not something that educational policy can directly affect and therefore we use measures of evenness when evaluating the effects of school performance information on school segregation.

We follow Massey and Denton (1988) and report the Duncan Dissimilarity Index. The Dissimilarity Index (DDI) is defined as $DDI = \frac{1}{2}\sum_{i=1}^{n} |\frac{a_i}{A} - \frac{b_i}{B}|$ where a_i is the number of individuals in group A in school i and b_i is the number of individuals in group B in school i, and A and B measure the total number of individuals in these two groups. The Dissimilarity Index ranges from zero (no segregation) to one (total segregation) and has a clear interpretation as the percentage of one of the two groups that would have to move to a different organizational unit (school), in order to produce a distribution in each organizational unit that matches the distribution of the entire population (zero segregation) (Duncan and Duncan, 1955).

The Duncan Dissimilarity index is however not without drawbacks. Most notably, it fails the transfer principal (James and Taeu-

ber (1985) and White (1986)). This means that it is insensitive to redistribution of minority group members among organizational units with minority proportions above or below the overall minority proportion. Only transfers of minority members from units where they are overrepresented to units where they are underrepresented (or vice versa) affect the value of the index. Therefore, we also report the effects on the Thiel Entorpy Index (Originally proposed by Theil and Finizza (1971)), another measure of evenness for which the transfer principle holds, in order not to miss any effects. The Thiel Entorpy Index (TEI) is defined as $TEI = \frac{1}{N}\sum_{i=1}^{N}\frac{x_j}{\mu}ln(\frac{x_j}{\mu})$ where N is the number of schools, x_j is a characteristic of school j (such as the share of students with a foreign background assigned to school j) and μ is the mean of x for all schools in N. The Thiel Entropy Index can be interpreted as the difference between the diversity (entropy) of the system and the weighted average diversity of individual organizational units, expressed as a fraction of the total diversity of the system (Reardon and Firebaugh, 2002).

Further, to determine the political feasibility of each priority structure, we look at the distance to the assigned school as well as whether the alternative priority structures manage to satisfy parents with regards to their school choices. We look at the fraction of students assigned to their preferred schools.

Finally, we abstract from factors that might be affected by the changing student composition. Firstly, each school's test score is assumed to be constant during the simulations, even though it may be affected by such changes. Secondly, we do not allow teachers to respond to the changes in student composition by taking a job at another school.

2.5.2 Preference Parameter Estimation

To estimate parental preferences, we need a model of how households behave in the school choice context. We assume that households are utility maximizing agents deriving utility from attending school j according to Equation (2.1). Assume that household i can choose between all schools $j \in J_i$, where J_i is the choice set of household i.[25] Let y_i be an indicator equal to j if household i chooses school j. Utility maximization requires that

$$y_i = j \quad \text{iff} \quad U_{ij} > U_{il} \quad \forall l \in J_i \tag{2.2}$$

The probability of household i choosing school j is then:

$$Pr[y_i = j] = Pr[U_{ij} > U_{il} \quad \forall l \in J_i] \tag{2.3}$$

Given our assumed distribution of ϵ_{ij} the choice probability can be expressed as:

$$Pr[y_i = j] = \frac{e^{\boldsymbol{X}'_{ij}\boldsymbol{\beta_i}}}{\sum_{k=1}^{m} e^{\boldsymbol{X}'_{ik}\boldsymbol{\beta_i}}} \tag{2.4}$$

The above equation can be estimated using a mixed logit model which relaxes the IIA assumption and thus avoids the restrictive substitution patterns IIA implies. Since $\boldsymbol{\beta}_i$ consists of random param-

[25] We define J as consisting of all schools in the municipality. Since the educational system is decentralized to the municipal level and households have the right to choose between any public school in their municipality, the natural definition of the choice set includes all schools located in the municipality. The Botkyrka School Choice Program provides information on and allows applications to all schools in Botkyrka, but not for schools located outside the municipality. The fact that we observe only three percent (see Table 2.4) choosing a school outside their municipality suggests that it is reasonable to restrict the choice set to schools located in the municipality. The limited number of schools ($N = 21$) also makes it reasonable to assume that households can consider all schools available in the municipality.

eters, we need to integrate the choice probability over all possible values of $\boldsymbol{\beta}$:

$$L_{ij}(\boldsymbol{\beta}_i) = Pr[y_h = j] = \int \frac{e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta_i}}}{\sum_{k=1}^{m} e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta_i}}} f(\boldsymbol{\beta}) d\boldsymbol{\beta} \tag{2.5}$$

where $f(\boldsymbol{\beta})$ is the density function of $\boldsymbol{\beta}$ (in this context also referred to as the mixing distribution).[26] As the coefficients are allowed to vary over individual households, we must make an assumption about their distribution. In our case, all variables are assumed to be random with a normal distribution, except the distance to school, if the school is in the same neighborhood as the student and whether a sibling is enrolled in grade K to 4 which are assumed to be fixed across individuals.

We have access to three ranked choices for each household. The fact that we have access to ranked data is helpful for identification. As Berry et al. (2004) point out, it gives us a "direct, data-based measure of substitution" which means that we can require the model to reproduce this observed pattern of substitution by restricting the parameters to match all three choices made by each household. Therefore, this is a source of variation that doesn't require any exogenous change in the choice sets. It is intuitive to think about this as three sequential choices. Let y_i^1 be the top choice, y_i^2 the second choice and y_i^3 the third choice. Then, the probability that household h would choose school j^1, j^2 and j^3 in that order is the product of the individual probabilities of these choices:

$$Pr[y_i^1 = j^1, y_i^2 = j^2, y_i^3 = j^3] = \int \prod_{r=1}^{3} \frac{e^{\boldsymbol{x}'_{ij^r}\boldsymbol{\beta_h}}}{\sum_{k \in J_i^r} e^{\boldsymbol{x}'_{ik}\boldsymbol{\beta}_i}} f(\boldsymbol{\beta}) d\boldsymbol{\beta} \tag{2.6}$$

[26] We specify a continuous mixing distribution. It is also possible to have a discrete mixing distribution, in which case the mixed logit model turns into a latent class logit model.

Note that the choice set is different for the top, second and third choice. For the first of the three terms in the product, the choice set J_i^1 consists of the $j = 1, ..., m$ schools. But once the household has listed their top choice, there are only $m - 1$ schools left to choose from. For the third choice, the choice set consists of $m - 2$ schools. Equation (2.6) describes the choice probability for a single household i. Summing over all households, we get the log likelihood function:

$$LL = \sum_i \sum_{j \in J^1} \sum_{j \in J^2} \sum_{j \in J^3} d_{ij}^1 d_{ij}^2 d_{ij}^2 ln(Pr[y_i^1 = j^1, y_i^2 = j^2, y_i^3 = j^3]) \tag{2.7}$$

where d_{hj}^r is equal to one if household i listed school j on rank r. In this setup there is no closed form $\boldsymbol{\beta}$. Therefore, simulation is used to obtain draws of $\boldsymbol{\beta}$ from its distribution $f(\boldsymbol{\beta})$ and for each draw, calculating the choice probability in equation 2.6. This is repeated 200 times, and inserting the simulated choice probability in equation (2.7) gives us an estimated log likelihood. The value of the parameters that maximize this equation is the maximum simulated likelihood estimator, or in other words, the value of the parameters of the distribution of $\boldsymbol{\beta}$ that maximizes the likelihood of us observing the school choices made by households.

The results are presented in Table 2.5.[27] We estimate the model on the entire population (pooling all cohorts 2011 to 2014) as well as for subgroups defined by foreign background and parental education. Estimation by subgroups will become important for our simulations, since Table 2.5 suggests that preferences vary by group. The school

[27]In general, the results are in line with previous papers on parental preferences for school attributes, suggesting that parents value proximity and peer composition. Our results deviate in that school performance does not seem to be as important as previous studies suggest. There are several potential explanations for this difference. For instance, the fact thatwe are studying school choice for younger children compared to most previous studies.

attributes included in the model are measured one year prior to the school choices being made. The coefficients should be interpreted as the change in utility from a one unit change in the corresponding school attribute. Standard errors are clustered at the 250×250 meters level (Table 2.11 presents results with clustering at the 500×500 and 750×750 level as well. Changing the level of clustering does not affect the results in any significant way).[28]

[28]Note that we are abstracting from modelling choice of residential location. Since admission in the Botkyrka School Choice Program is dependent upon one's residential location relative to school locations, households may choose their residential location in order to increase their probability of acceptance at a preferred school. If this behavior is present, we could overestimate the preferences for proximity relative to other school attributes. To test whether this is an issue in our setting, we assume that households would list the school they moved close to, as their top choice. We estimate our preferred specification on the whole population excluding the top choices. In addition, we estimate the same model on the subset of households that have not changed their residential location since their child was born - reasoning that these households are unlikely to have chosen their residential location based on the location of a specific primary school. Finally, we reestimate our model with a choice set restricted to the 10 schools closest to a student's home. The results from these exercises are presented in Table 2.12 in Appendix E. The results do not differ much from those in our preferred specification, with the exception that the preference parameter for enrolling your child in the same school as a sibling, drops when top choices are excluded. This is expected, since excluding the top choice is likely to be equivalent to excluding the school where siblings already are enrolled, from their choice set. We conclude that residential sorting does not seem be a significant problem in our setting. Also note that the variables included in our main model are motivated by what previous research has found to matter when households are choosing schools. As we have a rich data set including a multitude of school attributes, we can however test whether inclusion of additional variables matter. Table 2.10 in Appendix E presents the results from our extended model including a broader set of school attributes (definitions of these variables can be found in Table 2.7 in Appendix A) including a dummy for voucher schools, whether the school offers all grades up to grade 9, average test scores for students of their own type, the fraction of certified teachers, the student-teacher ratio, whether their own ethnic group is overrepresented at the school, the share of students from their own socioeconomic group, the share of newly arrived students at the school, the average income level

Note that we are not claiming that these parameters represent the true underlying preferences for school attributes but rather that we have put enough structure on this problem to be able to use the parameters in our counterfactual analysis. To further convince ourselves that the estimated preference parameters are valid for counterfactual analysis, we perform a validation exercise. First, we randomly divide our sample into five equally sized folds. We reestimate our model five times, each time excluding one of the folds and use the estimated parameters to construct complete rankings for those in the excluded fold. We then compare our predicted school rankings to the observed lists. Figure 2.3 displays histograms of the absolute difference between our predicted rankings and the choices observed i the data, separately for the top, second and third choices. We note that we can accurately predict the top choice of 73 percent. 42 percent of second choices are accurately predicted and 24 percent of third choices. We also note that few predictions are very inaccurate. For example, for the top choice our prediction is within two ranks for 89 percent of the population. Corresponding numbers for second and third choices are 82 and 78 percent.

2.5.3 Assessment of Truthfulness

This section is devoted to assessing the truthfulness in the households' submitted choice lists. We have shown that our estimated preferences predict choices well in our observed data. This is however not enough for our counterfactual analysis to be valid. We need our estimated preferences to predict choice behavior not just in the setting they were estimated but also under altered priority structures. For this we need two things, (i) no strategic incentives in the counterfactual

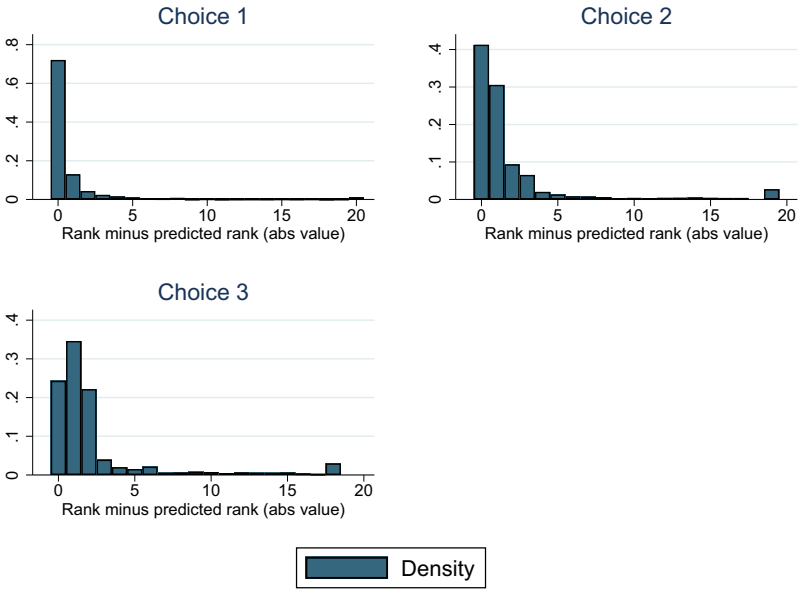
and unemployment events of parents and the share of households receiving social benefits. The inclusion of these variables does not substantially change the coefficients of the variables included in our preferred specification.

Table 2.5: Preference parameters from rank ordered mixed logit model

	All		Native				Foreign			
			High educ		Low educ		High educ		Low educ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Home-school distance	-1.788***		-2.053***		-1.775***		-1.535***		-1.809***	
	(0.0626)		(0.108)		(0.0763)		(0.0954)		(0.0850)	
Sibling in school	3.085***		3.464***		3.127***		3.131***		2.769***	
	(0.0922)		(0.235)		(0.190)		(0.230)		(0.143)	
Neighbourhood school	2.579***		2.770***		2.565***		1.871***		1.933***	
	(0.129)		(0.231)		(0.199)		(0.235)		(0.256)	
School test score	0.0393	0.109	0.102**	-0.201	0.0471	-0.179**	0.0603	0.135	-0.00828	0.246***
	(0.0302)	(0.119)	(0.0518)	(0.124)	(0.0415)	(0.0903)	(0.0612)	(0.644)	(0.0513)	(0.0854)
Foreign (share)	-2.677***	2.757***	-4.155***	-2.636***	-4.290***	-2.897***	-1.771***	2.521***	-0.900**	-2.761***
	(0.241)	(0.229)	(0.376)	(0.678)	(0.337)	(0.331)	(0.421)	(0.290)	(0.368)	(0.453)
Highly-educated (share)	3.024***	0.765	4.826***	1.072	2.189***	-0.860	1.360**	0.918	1.223**	0.102
	(0.316)	(0.581)	(0.481)	(2.698)	(0.474)	(0.737)	(0.557)	(2.074)	(0.607)	(0.596)
Observations	206405		70180		51008		37655		47562	

Note: This table presents the estimated mean and standard deviation of the preference parameters. The sample includes all primary school starters in Botkyrka 2011-2014. *Home-school distance* is the logarithm of the distance between the home and the school, measured in log(meters). *Sibling in school* is a dummy for having a sibling in the school. *Neighborhood school* is a dummy for the school being located in the student's own neighborhood. *Test scores* is the (standardized) score on the national exams taken by all students in all schools in year 3. *Foreign* is the fraction of students at the school that are (i) born abroad or (ii) have both parents born abroad and *Highly-educated* is the fraction of students with parents that have at least some post-secondary education. Standard errors are clustered at the 250 times 250 meters level.

Figure 2.3: Cross validation



Note: This figure displays histograms of our predicted rank of choices for top, second and third choices separately using a 5-fold cross validation approach.

scenarios and, (ii) truth-telling households in our observed data. As explained in section 2.4 our three counterfactual scenarios should not give rise to any strategic incentives and therefore the first point is not a problem. When it comes to truth-telling in our observed data there are two concerns, truncated lists and the possibility that households are "skipping the impossible".

We start with the problem of truncated lists. Although deferred acceptance is strategy proof, Haeringer and Klijn (2009) show that a restriction on households regarding how many schools they can rank reintroduces strategic incentives. Households may want to include one (or more) "safe school(s)" to avoid a situation where their child is not admitted to any of their submitted choices. It is not uncommon for school choice programs to have a limitation on the number of schools allowed on the application list. Abdulkadirolu et al. (2005a) and Abdulkadirolu et al. (2009) report that in NYC, students applying to high schools were allowed to list five programs before the redesign of the system in 2003-2004 and twelve programs afterwards, with more than 500 programs to choose from. With the five-school restriction, about 50 percent of households ranked less than the maximum allowed number of schools, after the redesign the corresponding number was 72-80 percent. Hence, the restricted number of choices does not necessarily induce strategic incentives for all households. Nevertheless, we will assess the application lists using the concept of "safe schools" to say something about the degree of truthfulness (during the discussion of truncated lists we will abstract from the problem of "skipping the impossible") in our setting.[29]

Let school j be defined as a safe school for household i if the probability of admission to this school is equal to 1. Then, in our setting with three choices, we claim the following.

Claim 1 *If the top choice is a safe school, the top choice is truthful.*

[29]This is based on Calsamiglia et al. (2010).

Claim 2 *If the top choice is not a safe school but the second choice is a safe school, both the top and second choice are truthful.*

Proofs of these claims are provided in Appendix C, together with the necessary assumptions. These claims imply that the submitted choice lists can be considered truth-revealing if they include a safe choice, unless the first safe choice is also the last choice.[30]

We take this concept to the data. We have access to the school capacity of the public schools that the municipality used in their placement algorithm for the years 2012-2014. We start by defining all public schools that were undersubscribed in all these three years as safe schools for all students. We define an undersubscribed school as a school that has more capacity than students assigned to it once the term starts. Even though we lack school capacity data for 2011 we claim that this is a fair definition as the cohort of school starters in 2011 is, by some margin, the smallest cohort of the four cohorts we observe. Next we define school j as safe for household i if all households observed during 2011 to 2014 residing in the same geographic unit (250 $\times$ 250 m) who applied to school j were admitted or if household i already has a child enrolled in school j in grade K to 4. Note that we cannot determine whether a voucher school is a safe school or not.[31] Using these definitions, we find that more than 95 percent of all households list a safe school as either their top or second choice. This is due to the previously mentioned, significant overcapacity that the Borkyrka school choice program operates under. Note also that the remaining 5 percent are not categorized as non-truthful, rather, they we can not categorize them as truthful or not. Given this, we claim that non-truthfulness due to truncated lists

[30] Note that this generalizes to truncated lists of any length.

[31] As mentioned earlier, voucher schools admit students on a first-come first-served basis and we do not observe the time spent in the school's queues. Nor do we have access to the actual capacity of the voucher schools.

is not a problem in this setting.

Next we have the problem of "skipping the impossible". This problem arises because the DA mechanisms truth-telling properties rely on week dominance (Fack et al., 2017). This means that the households do not really have an incentive to reveal the truth but rather they have no incentive to not do so. Hence, the mechanism could give rise to multiple equilibria. The main implication of this is that households can be inclined to "skip the impossible", i.e. not list schools that they perceive as impossible to be admitted to. As changes to the priority structures can alter the set of schools that are "impossible" for different groups of students to get admitted to, this is a concern. Again, the overcapacity in Botkyrka is helpful to us. Between 2012 and 2014 not one of the 18 public schools were oversubscribed for all three years. One school was oversubscribed two out of three years and two where oversubscribed one of three years. Based on this we argue that the problem with "skipping the impossible" is limited. During the relevant time period there are, in reality, no schools that are impossible for any student to get into.[32] In our counterfactual scenarios situations could arise where specific schools are "impossible" to get into for certain students. Especially as we increase the cohort size. This is however not a problem as the outcomes of interest in our counterfactual scenarios are related to assignment and, by definition, assignment is not affected by the household's choice whether to include "impossible schools" or not.

Taken together, the significant overcapacity in Botkyrka during this time period helps us mitigate the problem of households not being truth-telling. In fact, given the overcapacity in Botkyrka, our estimation, based on truthfulness, would be almost identical if we

[32]We cannot tell if any of the five voucher schools should be defined as "impossible", however, given that four of them are small and cater to specific groups we do not believe that this is of much importance.

were to estimate based on stability (as recommended in Fack et al. (2017)). The main difference would be that while estimating based on stability, we would have restricted our focus to only the students' top choices.

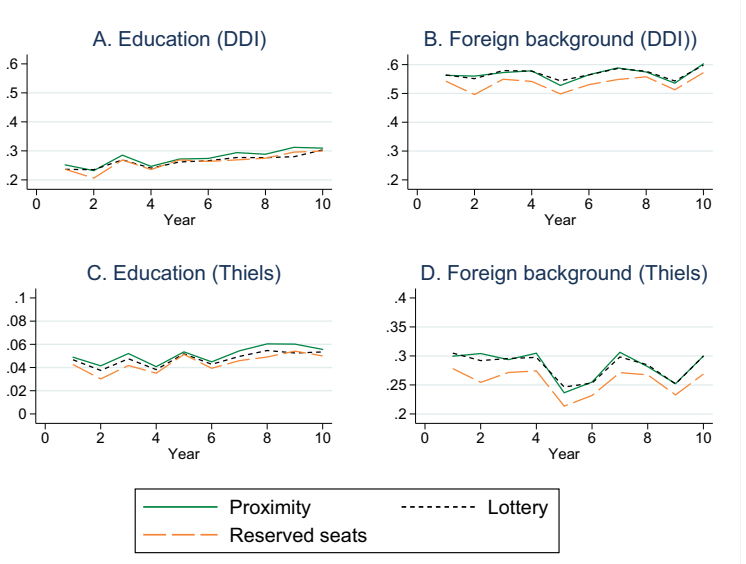
2.6 Results

Figure 2.4 displays the level of segregation in terms of parental education and foreign background for each cohort entering the school system. We note that the evolution of the segregation over time does not seem to vary with the priority structure implemented.[33] The lack of differences in time trends between the counterfactual scenarios studied suggests that the initial reallocation of students is not large enough to incentivize the next cohort of school choosing households to alter their choice behavior. This allows us to pool the simulations over the years and look at the distribution of the simulation outcomes. These distributions can be seen in Figure 2.5. When it comes to segregation in terms of foreign background, the distribution clearly shifts to the left when we go from a proximity-based priority structure to one based on reserved seats. The DDI decreases with 4 points (from 0.57 to 0.53) and the Theil Index decreases with 2 points (from 0.28 to 0.26). We can see a similar pattern in segregation in terms of parental education. When going from a priority structure based on proximity to one based on reserved seats the DDI decreases with 2 points (from 0.28 to 0.26) and the Theil Index decreases with 1 point (from 0.05 to 0.04). The lottery-based priority structure produces similar levels of school segregation as the priority structure based on proximity.

For school districts considering whether to implement a new priority structure, an important concern may be the expected response of

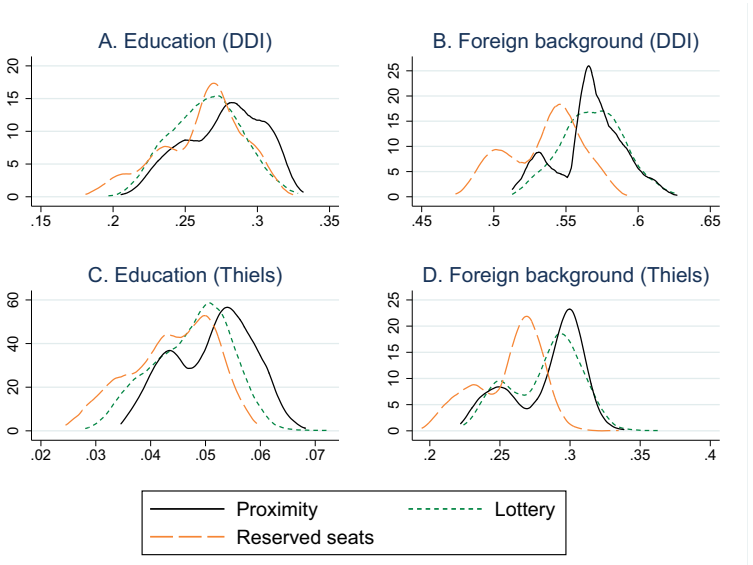
[33]This is true for all outcomes of interest. This can be seen in Figure 2.12.

Figure 2.4: Level of segregation under alternative priority structures over time (mean)



parents and students living in their district. For this reason, we study outcomes related to student welfare. Firstly, when looking at the primary school level, proximity of the school is an important consideration. Secondly, the legitimacy of the school choice system may depend on whether or not it is able to satisfy households' choices. Subfigure A in Figure 2.6 presents the distribution of simulations results for the share of students assigned their most preferred school. We can see that the the priority structure based on proximity performs best in this dimension with 83 percent of all households assigned their most preferred school. It is however only outperforming the priority structure based on reserved seats with less than one percentage point. The lottery-based priority structure performs a little worse, assigning 82 percent of households to their most preferred school. In Subfigure

Figure 2.5: Level of segregation under alternative priority structures (kernel density functions)



B we look at the share of students assigned to any of their three most preferred schools. The proximity-based priority structure and the priority structure based on reserved seats, perform equally well in this dimension, assigning 90 percent of students to one of their top three most preferred schools. The priority structure based on lottery performs a little less well than the other systems, assigning 89 percent of students to one of their top three preferred schools. Subfigure C in Figure 2.6 shows the simulation results for the average distance from the students home to their assigned school for the different priority structures. In terms of distance to the assigned school, we observe an increase of about 40 meters (from 957 to 997) when abandoning a proximity-based priority structure and instead implementing one based on reserved seats. Again, the lottery-based priority structure

Figure 2.6: Student welfare (kernel density functions)

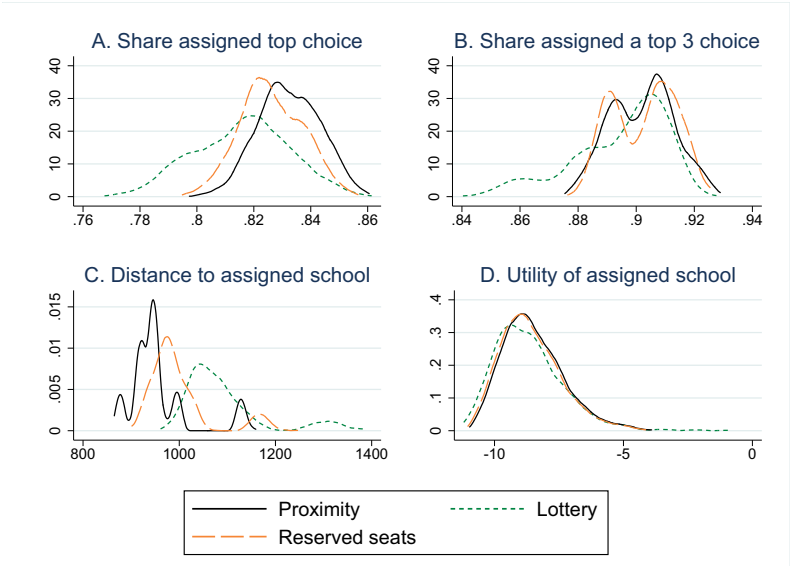
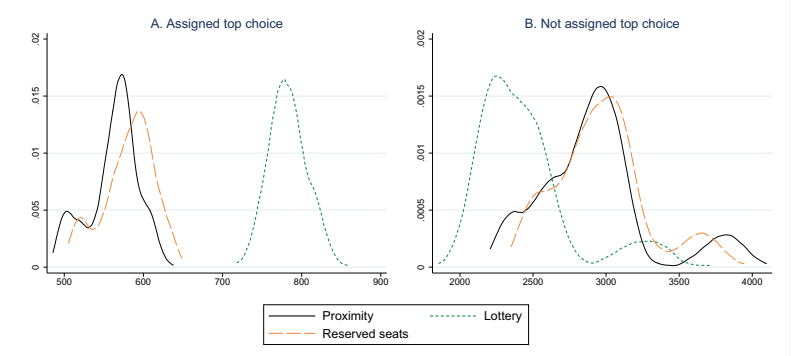


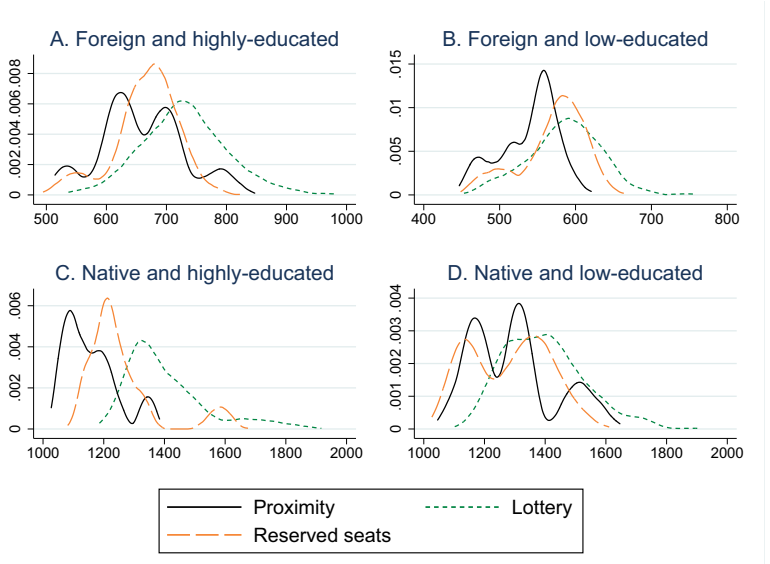
Figure 2.7: Home-school distance by assigned top choice or not (kernel density functions)



preforms worst with an average distance to school of 1,090 meters. In Subfigure D in Figure 2.6 we also present the utility. The differences

between the different priority structures are small.[34]

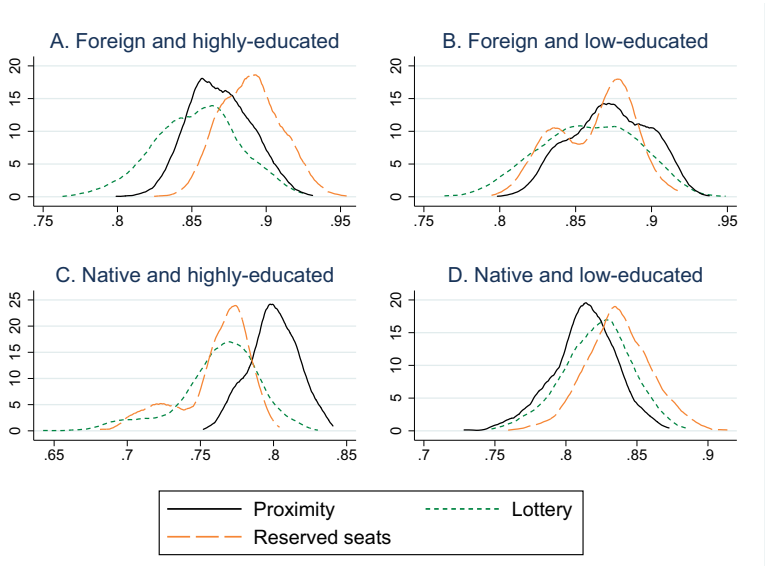
Figure 2.8: Home-school distance by subgroup (kernel density functions)



In Figure 2.7 we present results for the average distance to school but with the sample divided into the households that were admitted to their most preferred school and those that were not. We can see that the increase in distance, when going from a proximity-based priority structure to a priority structure based on reserved seats that we observed in Figure 2.6, is driven by those assigned to their most preferred school. The students not assigned to their most preferred school do not end up in schools further away from home. This implies

[34]Given the ordinal nature of utility and the fact that it is unclear to what extent our estimated parameters represent true underlying preferences, we will not be drawing any conclusions about the utility. We do however include it for completeness.

Figure 2.9: Share admitted to top choice by subgroup (kernel density functions)



that this increase in distance is driven by students that want to travel further in order to get into a more preferred school. Therefore it is unclear to what extent it should be considered a cost. We can also see that the lottery-based priority structure deviates considerably from the two other systems with a longer distance for those assigned to their most preferred school and a shorter distance for those who are not. This is most likely due to this priority structure allowing students from the South of the municipality to get into their most preferred schools, which the two other priority structures will not.

In Figure 2.8 and 2.9 we show how the proxies for welfare vary for different subgroups depending on the priority structure implemented. We can see that both in terms of distance to school and the probability of being admitted to one's most preferred school, highly-

educated natives are best off when priority is based on proximity. Highly-educated and foreign households and low-educated and native households, instead, benefit most from the priority structure based on reserved seats.

2.7 Concluding Remarks

In this paper we study the effects on school segregation and student welfare when altering the priority structure in a school choice program using deferred acceptance to assign students to schools. Three types of priority structures are evaluated: proximity-based, lottery-based and one based on reserved seats. Given the widespread use of the deferred acceptance mechanism in school choice programs and the potential importance of priority structures when allocating students, this is an understudied subject.

We find that segregation, both in terms of foreign background and parental education, is lower when using a priority structure based on reserved seats compared to when using a proximity- or lottery-based priority structure. The reallocation of students is immediate after implementation of the new priority structure and is stable over time. When compared to a system based on proximity, this reduction in segregation comes with a small increase in the average distance to school and a minor decrease in the share of students admitted to their most preferred school. The increase in average distance to school is however driven by students being admitted to their most preferred schools. The lottery-based system underperforms compared to the two other systems in these dimensions. Even though the system based on proximity and the one based on reserved seats produces similar results on average, there are clear winners and losers from the two systems. Highly-educated native households lose their residential privilege when proximity-based priorities are abandoned, while low-

educated households with a native background and highly-educated households with a foreign background benefit from a system based on reserved seats.

Botkyrka is a very segregated municipality with distinct neighborhoods clearly separated from each other. This may affect the extent to which our results generalize to other settings. The effect on segregation implementing a priority structure based on reserved seats or a lottery may be muted because most parents do not consider it to be an option to allow their six-year old to commute to another neighborhood several kilometers away. In other settings and with older children the effects are likely to be larger. With that said, the observed effects on school segregation when reserving seats are not large, implying that it might not be the silver bullet we can use to eradicate school segregation. Still, the effect on school segregation is not negligible and the cost in terms of satisfying households' preferences for schools and the additional distance to assigned schools are small. Reserved seats for underrepresented groups should therefore not be discarded as a tool to reduce school segregation. When considering changing from a proximity-based priority structure to one based on reserved seats one should however be clear that it comes with a redistribution of welfare. Further, the results seem to imply that a system with lottery-based priorities does not perform well in any dimension compared to the two other systems.

Bibliography

Abdulkadirolu, A., Pathak, P. A., and Roth, A. E. (2005a). The New York City High School Match. *AEA Papers and Proceedings*, 95(2):364–367.

Abdulkadirolu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match. *American Economic Review*, 99(5):1954–1978.

Abdulkadirolu, A., Pathak, P. A., Roth, A. E., and Sönmez, T. (2005b). The Boston Public School Match. *AEA Papers and Proceedings*, 95(2):368–371.

Abdulkadirolu, A. and Sönmez, T. (2003). School Choice: A Mechanism Design Approach. *American Economic Review*, 93(3):729–747.

Allen, R. and Vignoles, A. (2007). What should an index of school segregation measure? *Oxford Review of Education*, 33(5):643–668.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy*, 112(1):68–105.

Bifulco, R. and Ladd, H. F. (2007). School Choice, Racial Segregation, and Test-Score Gaps: Evidence from North Carolina's Charter School Program. *Journal of Policy Analysis and Management*, 26(1):31–56.

Billings, S., Deming, D. J., and Ross, S. L. (2016). Partners in Crime: Schools, Neighbourhoods and the Formation of Criminal Networks. *NBER Working Paper, No. 21962.*

Billings, S. B., Deming, D. J., and Rockoff, J. (2013). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg. *Quaterly Journal of Economics*, 1(129):435–476.

Böhlmark, A., Holmlund, H., and Lindahl, M. (2016). Parental Choice, Neighborhood Segregation or Cream Skimming? An Analysis of School Segregation after a Generalized Choice Reform. *Journal of Population Economics*, 29(29):1155–1190.

Böhlmark, A. and Lindahl, M. (2015). Independent Schools and Long-run Educational Outcomes - Evidence from Sweden's Large Scale Voucher Reform. *Economica*, 82(327):508?551.

Borghans, L., Golsteyn, B. H., and Zölitz, U. (2015). Parental Preferences for Primary School Characteristics. *The B.E. Journal of Economic Analysis and Policy*, 15(1):1?33.

Burgess, S., Greaves, E., Vignoles, A., and Wilson, D. (2015). What Parents Want: School Preferences and School Choice. *The Economic Journal*, 125(587):1262?1289.

Calsamiglia, C. and Güell, M. (2014). The Illusion of School Choice: Empirical Evidence from Barcelona. *Barcelona Graduate School of Economics Working Paper, No. 810*.

Calsamiglia, C., Haeringer, G., and Klijn, F. (2010). Constrained School Choice: An Experimental Study. *American Economic Review*, 100(4):1860–1874.

De Haan, M., Gautier, P. A., Oosterbeek, H., and Van der Klaauw, B. (2015). The Performance of School Assignment Mechanisms in Practice. *IZA Discussion Paper, No. 9118*.

Duncan, O. D. and Duncan, B. (1955). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, 20(2):210–217.

Ernst & Young AB (2014). Kan resursfördelningen lösa skolkrisen? En studie om resursfördelning till grundskolan.

Fack, G., Grenet, J., and He, Y. (2017). Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions. *PSE Working Paper, No. 2015-35*, 1(1):1–11.

Gale, D. and Shapley, L. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9–15.

Gallego, F. A. and Hernando, A. E. (2008). On the Determinants and Implications of School Choice: Semi-Structural Simulations for Chile (with comments). *Economia*, 9(1):197–244.

Gamoran, A. and An, B. P. (2016). Effects of School Segregation and School Resources in a Changing Policy Context. *Educational Evaluation and Policy Analysis*, 38(1):43–64.

Haeringer, G. and Klijn, F. (2009). Constrained School Choice. *Journal of Economic Theory*, 144(5):1921–1947.

Hafalir, I. E., Yenmez, M. B., and Yildirim, M. A. (2013). Effective Affirmative Action in School Choice. *Theoretical Economics*, 0(8):325–363.

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2000). How Much Does School Integration Affect Student Achievement? *Paper prepared for the Annual Meetings of the Association for Public Policy Analysis and Management*.

Hastings, J. S., Kane, T. J., and Staiger, D. O. (2009). Heterogeneous Preferences and the Efficacy of Public School Choice. *NBER Working Paper, No. 2145*.

He, Y. (2017). Gaming the Boston School Choice Mechanism in Beijing. *TSE Working Paper, No. 15-551*.

Hsieh, C.-T. and Urquiola, M. (2006). The Effects of Generalized School Choice on Achievement and Stratification: Evidence from Chile's School Voucher Program. *Journal of Public Economics*, 90(8-9):1477–1503.

James, D. R. and Taeuber, K. E. (1985). Measures of Segregation. *Sociological Methodology*, 15:1–32.

Johson, R. C. (2015). Long-run Impacts of School Desegregation and School Quality on Adult Attainments. *NBER Working Paper, No. 16664*.

Kojima, F. (2012). School Choice: Impossibilities for Affirmative Action. *Games and Economic Behavior*, 75:685–693.

Massey, D. S. and Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, 67(2):281–315.

Pathak, P. A. and Shi, P. (2017). How Well Do Structural Demand Models Work? Counterfactual Predictions in School Choice. *NBER Working Paper, No. 24017*.

Reardon, S. F. and Firebaugh, G. (2002). Measures of MultiGroup Segregation. *Sociological Methodology*, 31(1):33–67.

Söderström, M. and Uusitalo, R. (2010). School Choice and Segregation: Evidence from an Admission Reform. *The Scandinavian Journal of Economics*, 112(1):55–76.

Theil, H. and Finizza, A. J. (1971). A note on the measurement of racial integration of schools by means of informational concepts. *The Journal of Mathematical Sociology*, 1(2):187–193.

White, M. J. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index*, 52(2):198–221.

Appendix A - Variable Definitions

Table 2.6: Definition of variables at the individual and individual-school level

Variable	Definition
Individual level	
Home school distance	Distance between the students residential location and the school location. Measured in meters for the descriptive tables and in log of meters in the mixed logit models.
Sibling in the school system	Dummy equal to 1 for primary school starters with a sibling that will attend grade 1 to 9 in any school in the municipality.
Number of schools in nbd	Number of schools located in the neighbourhood (Norra Botkyrka, Tumba, Tullinge) in which the student lives.
Disposable income	Average yearly parental disposable income measured in 100SEK.
Highly educated	A dummy equal to 1 if at least one parent have studied at a level higher than upper secondary school. At the school level, we measure the fraction of students with highly educated parents.
Foreign background	A dummy equal to 1 if the student is born abroad or if both parents are born abroad. At the school level, we measure the fraction of students with foreign background.
Impoverished	A dummy variable equal to 1 if the household received any social assistance or housing allowances. At the school level, we measure the fraction of students from impoverished households.
Days unemployed	Parental average number of days with unemployment insurance during the year.
Individual-school level	
Sibling in school	Dummy equal to 1 for primary school starters with a sibling that will attend the school applied to in grade 1 to 9.
Neighbourhood school	Dummy variable equal to 1 if the school is located in the same neighbourhood as the student's residential location.
Test scores own group	Measures for each individual and each school the average test score for students with the same immigration and educational background.
Own ethnicity overrepresented	Measures for each individual and each school whether the students ethnic group is overrepresented or not (i.e. the share of the ethnic group at the school is larger than the share of that ethnic group in the municipality).
Share own socioeconomic group	Measures for each individual and each school the share of students in the same socioeconomic group.

Table 2.7: Definition of variables at the school level

Variable	Definition
Pass rate	Fraction of students passing the standardized tests in year 3.
Certified teachers	Fraction of certified teachers at the school.
Student-teacher ratio	Number of students at the school divided by the number of teachers at the school.
Student well-being	The schools average score on the student-surveys administered by the municipality to all students enrolled in a public school, asking about the students well-being at the school.
Foreign background	Share of students at the school with foreign background, defined as being born abroad or having both parents born abroad.
Highly educated	A student is classified as having highly educated parents if at least one parent have at lest 2 years of education at a level higher than upper secondary school.
Average school capacity	The average of the number of seats available for primary school starters at each school.
Newly arrived	Share of students at the school that are newly arrived, i.e. arrived to Sweden less than 4 years ago.
Average disposable income	Parental yearly average income for each student, where income is measured in 100SEK.
Share impoversihed	The share of students at the school classified as impoverished.
Average days unemployed	The parental average number of days as unemployed during the preceding calendar year.
Socioeconomic composition	The school's average score on the socioeconomic index, as defined in table 2.6.
All grades up to 9	Dummy variable equal to one if the school offers all grades from 0 to 9.
Boys	Male share of the students at the school.
Voucher school	Dummy variable equal to one for voucher school and zero otherwise. Voucher schools are publicly funded, but privately operated.
Test score	Average test score of students on the standardized tests in year 3, at the school level.
Student well-being	The schools average score on the student-surveys administered by the municipality to all students enrolled in a public school, asking about the students well-being at the school.

Appendix B - Additional Summary Statistics

Table 2.8: Summary statistics of primary school starters not admitted at their top choice school

	Total mean/sd	North mean/sd	West mean/sd	East mean/sd
Native, highly educated	0.23	0.11	0.28	0.65
	(0.42)	(0.31)	(0.45)	(0.48)
Native, low educated	0.20	0.16	0.29	0.25
	(0.40)	(0.36)	(0.46)	(0.44)
Foreign, highly educated	0.26	0.31	0.25	0.05
	(0.44)	(0.46)	(0.44)	(0.23)
Foreign, low educated	0.32	0.43	0.18	0.04
	(0.47)	(0.50)	(0.39)	(0.19)
Sibling in the school system	0.27	0.27	0.29	0.27
	(0.45)	(0.44)	(0.46)	(0.45)
Distance to attended school	1541.16	1596.27	1798.17	1038.31
	(1685.24)	(1875.04)	(1642.25)	(872.73)
Distance to top choice	1796.12	1988.28	2114.89	896.50
	(1747.19)	(1973.18)	(1537.05)	(884.80)
Observations	367	236	76	55

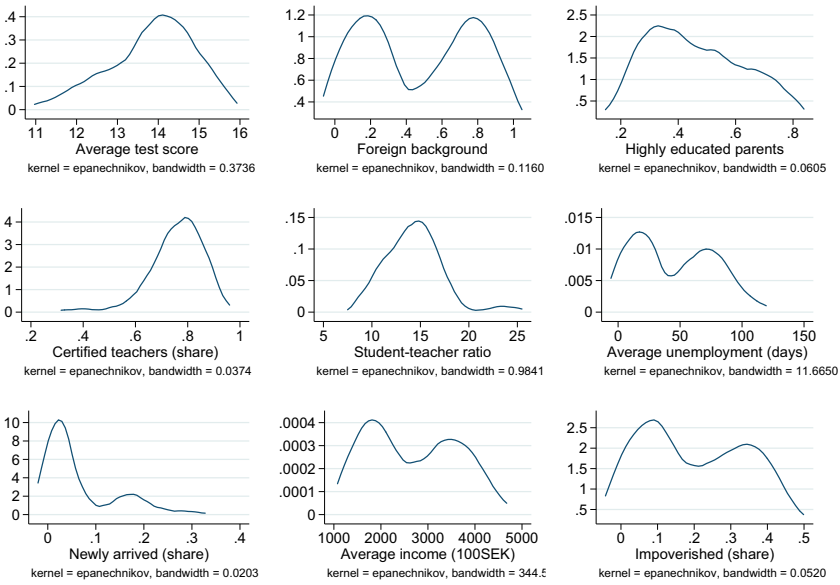
Note: This table presents the mean and standard deviation of each variable. The sample includes all primary school starters choosing schools in Botkyrka 2011-2014, that were not admitted at their top choice school. *Highly-educated* indicates that at least one parents have an education level above upper secondary school. *Foreign* indicates that the student was born abroad or has both parents born abroad, otherwise the student is categorized as *native*. *Sibling in the school system* indicates if there is a sibling currently attending grade K-9 in any primary school in Botkyrka. *Distance to attended school* and *top choice* measures the home-school distance to the attended school and the school listed as the top choice in meters.

Table 2.9: Extended summary statistics of schools

	Total	North	West	East
	mean/sd	mean/sd	mean/sd	mean/sd
Average test score	13.84	13.27	14.03	14.59
	(1.00)	(0.94)	(0.91)	(0.57)
Foreign background	0.48	0.80	0.32	0.12
	(0.31)	(0.09)	(0.18)	(0.03)
Highly educated parents	0.46	0.32	0.49	0.67
	(0.16)	(0.07)	(0.10)	(0.09)
School capacity	54.94	55.00	55.09	54.60
	(18.15)	(12.62)	(24.15)	(17.97)
Voucher school	0.10	0.11	0.14	0.00
	(0.30)	(0.32)	(0.36)	(0.00)
Certified teachers (share)	0.76	0.74	0.74	0.81
	(0.10)	(0.08)	(0.13)	(0.08)
Student-teacher ratio	14.29	12.66	14.54	16.89
	(2.93)	(1.75)	(3.05)	(2.52)
Newly arrived (share)	0.06	0.11	0.04	0.01
	(0.07)	(0.08)	(0.05)	(0.01)
Average income (100SEK)	2681.17	1799.38	3003.58	3817.00
	(928.64)	(276.64)	(588.13)	(387.67)
Impoverished (share)	0.20	0.31	0.17	0.05
	(0.14)	(0.09)	(0.12)	(0.03)
Average unemployment (days)	45.30	76.45	29.37	11.54
	(31.44)	(15.35)	(17.95)	(3.70)
Offers all grades up to 9	0.67	0.78	0.86	0.20
	(0.47)	(0.42)	(0.36)	(0.41)
Observations	84	36	28	20

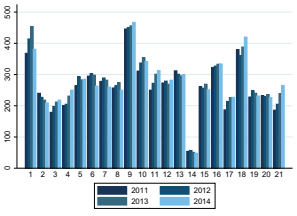
Note: This table presents the mean and standard deviation of each variable. The sample includes one observation per year of all schools available for choice in Botkyrka 2011-2014. See Appendix A for variable definitions.

Figure 2.10: Kernel density estimates of school attributes

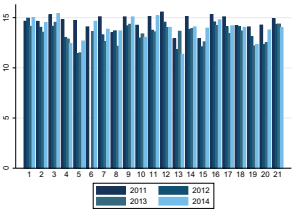


Note: These graphs are produced using the pooled sample for 2011-2014. *Average test scores* is the school average test score (max = 20) on the standardized tests in Swedish and mathematics taken in grade 3. *Foreign background* equals one for students born abroad or with both parents born abroad. *Highly-educated* indicates the share of students with at least one parent with higher education (above upper secondary school). *Certified teachers* is the share of certified teachers at the school. *Student-teacher ratio* indicates the number of students per teacher. *Average unemployment days* indicates the parental yearly average of days with unemployment insurance. *Newly arrived* indicates the share of students that arrived to Sweden less than 4 years ago. *Average income* is the average parental income measured in 100SEK. *Impoverished* indicates the share of students living in households that receive social assistance or housing allowances.

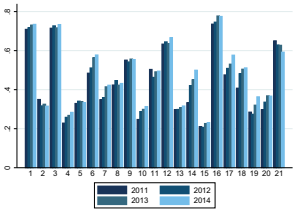
Figure 2.11: School attributes over time



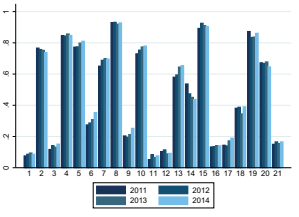
(a) Number of students (K-5)



(b) Average test score



(c) Share highly-educated parents



(d) Share foreign background

Note: This graph presents the *number of students enrolled in grade K-5*, *the average test score* (max = 20) on the standardized tests in Swedish and mathematics taken in grade 3, *the share with highly-educated parents* and *the share of students with foreign background* by year (2011-2014) for each school numbered from 1 to 21.

Appendix C - Truth-telling

Assumptions: Assume that all schools $j \in N$ are acceptable and that the households are not indifferent between any two schools $\in N$. Households are risk-neutral and rank the schools in order to maximize their utility as expressed in Equation 2.1. All schools have a strict priority ordering over students, independent of the households' ranking of schools. All students have a positive probability p of being accepted to any school $j \in N$, i.e. $\forall h, \forall j : p_j^i > 0$. p_j^i can be interpreted as household i's probability of having a priority above the capacity cut off for school j. A school j is defined as *safe* for household i if the probability of household h being admitted to school j is equal to 1, i.e. school j is a safe school for household h if $p_j^h = 1$. Assume also that households are able to determine whether a school is safe or not.

Proof of claim 1: Suppose that household h lists school j as their top choice, and that school j is a safe school. Since $p_j^h = 1$, the expected utility of submitting any ranking where school j is the top choice is U_j. Suppose now that there exist (at least) one other school k that is preferred to school j ($U_k > U_j$). If household h would modify the ranking and list school k as their top choice, their expected utility would be equal to $p_k U_k + (1 - p_k)U_j$. Since $0 < p_k < 1$ and $U_k > U_j$, it follows that $p_k U_k + (1 - p_k)U_j > U_j$ and household h would be better off with the modified ranking. Hence, if household h submits a safe school j as their top choice, there cannot exist any other school $k \in N$ that is preferred to school j.

Proof of claim 2: Suppose that household h lists the non-safe school k as their top choice and the safe school j as their second choice. The expected utility of this ranking is $p_k U_k + (1 - p_k)U_j$.

Suppose that school k is not preferred to school j ($U_k < U_j$). If household h would modify their ranking and list school j as their top choice, they would get utility U_j since $p_j = 1$. Since $U_k < U_j$ and $p_k > 0$, this implies that $U_j > p_k U_k + (1 - p_k)U_j$, which means that household h would be better off with the modified ranking. Hence, given that school j is ranked second, it must be that school k is preferred to school j. Next, suppose that there is another school l that is preferred to school j ($U_l > U_j$). If household h would modify their ranking and rank school l above school j (as their top or second choice), they would get expected utility $p_k U_k + p_l(1 - p_k)U_l + (1 - p_k - p_l(1 - p_k))U_j$. Since $U_l > U_j$ and $p_l > 0$, this implies that $p_k U_k + p_l(1-p_k)U_l + (1-p_k-p_l(1-p_k))U_j > p_k U_k + (1-p_k)U_j$, and household h would be better off with the modified ranking. Hence, given that school j is ranked second, there cannot exist any other school $l \neq k \in N$ that is preferred to school j.

Appendix D - Deferred Acceptance with Reserved Seats

Again, we assume that all students have strict preferences over all schools and that all schools have a strict priority ordering of all students. In addition, we introduce a finite number N of student types $j = 1, ..., N$. At each school, s_j seats are reserved for students of type j. The deferred acceptance algorithm modified to incorporate reserved seats would then assign students to schools in the following way:

Step 1 **a)** Each student applies to her top choice. Each school tentatively accepts one student at a time according to their priority ordering until there are no applicants left or the school run out of reserved seats for type j in which case all remaining applicants of type j are rejected.
b) If there is capacity left at the school (which happens when there are less applicants than reserved seats for at least one student type j), all students rejected in this round are considered together and the school tentatively accepts one student at a time according to their priority.

Step k **a)** Each student rejected in the previous round applies to her next preferred school. Each school considers all tentatively accepted and new applicants together and tentatively accepts one student at a time according to their priority ordering until there are no applicants left or the school run out of reserved seats for type j in which case all remaining applicants of type j are rejected.
b) If there is capacity left at the school (which happens when there are less applicants than reserved seats for at least one student type j), all students rejected in this round are considered together and the

school tentatively accepts one student at a time according to their priority.

The algorithm terminates when no student is rejected, at which point all students are placed at their final tentative assignment.

Appendix E - Alternative Specifications of the Mixed Logit Model

Table 2.10: Preference parameters from rank ordered mixed logit model: extended model with additional school attributes included

	Main model		Extended model	
	Mean	SD	Mean	SD
Home-school distance (log meters)	-1.788***		-1.959***	
	(0.0626)		(0.0584)	
Sibling in school	3.085***		3.243***	
	(0.0922)		(0.100)	
Neighbourhood school	2.579***		2.060***	
	(0.129)		(0.128)	
School average test score (stand)	0.0393	0.109	0.00484	0.0789
	(0.0302)	(0.119)	(0.0370)	(0.0616)
Foreign (share)	-2.677***	2.757***	-2.340***	2.076***
	(0.241)	(0.229)	(0.541)	(0.354)
Highly-educated (share)	3.024***	0.765	2.344***	-0.190
	(0.316)	(0.581)	(0.646)	(0.328)
Observations	206405		186395	
Additional controls	No	No	Yes	Yes

Note: This table present the estimated mean and standard deviation of the preference parameters. The sample includes all primary school starters in Botkyrka 2011-2014. *Home-school distance* is the logarithm of the distance between the home and the school, measured in log(meters). *Sibling in school* is a dummy for having a sibling in the school. *Neighborhood school* is a dummy for the school being located in the own neighborhood. *Test scores* is the (standardized) score on the national exams taken by all students in all schools in year 3. *Foreign* is the fraction of students at the school that are (i) born abroad or (ii) have both parents born abroad and *Highly educated* is the fraction of students with parents that have at least some post-secondary education. For additional variables included, see definitions in Table 2.6-2.7. Standard errors are clustered at the 250 times 250 meters grid.

Table 2.11: Preference parameters from rank ordered mixed logit model: varying the level of clustering

	Main model (250 m grid)		Main model (500 m grid)		Main model (750 m grid)	
	Mean	SD	Mean	SD	Mean	SD
Home-school distance (log meters)	-1.788***		-1.788***		-1.788***	
	(0.0626)		(0.0796)		(0.0870)	
Sibling in school	3.085***		3.085***		3.085***	
	(0.0922)		(0.103)		(0.122)	
Neighbourhood school	2.579***		2.579***		2.579***	
	(0.129)		(0.122)		(0.149)	
School average test score (stand)	0.0393	0.109	0.0393	0.109	0.0393	0.109
	(0.0302)	(0.119)	(0.0363)	(0.119)	(0.0450)	(0.137)
Foreign (share)	-2.677***	2.757***	-2.677***	2.757***	-2.677***	2.757***
	(0.241)	(0.229)	(0.299)	(0.260)	(0.308)	(0.291)
Highly-educated (share)	3.024***	0.765	3.024***	0.765	3.024***	0.765
	(0.316)	(0.581)	(0.479)	(0.681)	(0.617)	(0.743)
Observations	206405		206405		206405	

Note: This table present the estimated mean and standard deviation of the preference parameters. The sample includes all primary school starters in Botkyrka 2011-2014. *Home-school distance* is the logarithm of the distance (meters) between the home and the school. *Sibling in school* is a dummy for having a sibling in the school. *Neighborhood school* is a dummy for the school being located in the own neighborhood. *Test scores* is the (standardized) score on the national exams taken by all students in all schools in year 3. *Foreign* is the fraction of students at the school that are (i) born abroad or (ii) have both parents born abroad and *Highly-educated* is the fraction of students with parents that have at least some post-secondary education. Standard errors are clustered at the 250×250, 500×500 or 750×750 meters grid.

Table 2.12: Preference parameters from rank ordered mixed logit model: testing for residential sorting

	Total		No top choices		No movers		Restr. choice set	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Home-school distance	-1.788***		-1.608***		-1.874***		-1.778***	
	(0.0626)		(0.0777)		(0.0797)		(0.0666)	
Sibling in school	3.085***		1.704***		3.137***		3.139***	
	(0.0922)		(0.144)		(0.118)		(0.0886)	
Neighbourhood school	2.579***		2.785***		2.708***		1.722***	
	(0.129)		(0.129)		(0.181)		(0.138)	
Avgerage test score	0.0393	0.109	0.0280	0.0377	0.0593	-0.131	0.0347	0.160*
	(0.0302)	(0.119)	(0.0327)	(0.0603)	(0.0370)	(0.112)	(0.0308)	(0.0876)
Foreign (share)	-2.677***	2.757***	-2.537***	1.876***	-3.065***	2.841***	-3.504***	4.192***
	(0.241)	(0.229)	(0.217)	(0.307)	(0.321)	(0.290)	(0.340)	(0.339)
Highly-educated (share)	3.024***	0.765	3.340***	0.166	2.896***	0.934	3.537***	2.094***
	(0.316)	(0.581)	(0.310)	(0.126)	(0.393)	(0.678)	(0.319)	(0.691)
Observations	206405		134230		116207		93752	

Note: This table present the estimated mean and standard deviation of the preference parameters. The sample includes all primary school starters in Botkyrka 2011-2014. *Home-school distance* is the logarithm of the distance between the home and the school, measured in log(meters). *Sibling in school* is a dummy for having a sibling in the school. *Neighborhood school* is a dummy for the school being located in the own neighborhood. *Test scores* is the (standardized) score on the national exams taken by all students in all schools in year 3. *Foreign* is the fraction of students at the school that are (i) born abroad or (ii) have both parents born abroad and *Highly-educated* is the fraction of students with parents that have at least some post-secondary education. Standard errors are clustered at the 250 times 250 meters grid.

Appendix F - Additional Results

Figure 2.12: Level of student welfare under alternative priority structures over time (mean)

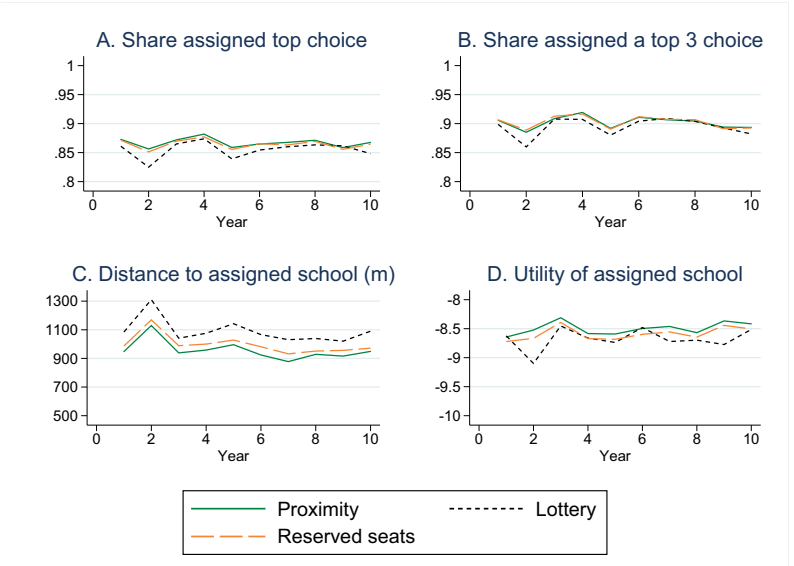


Figure 2.13: Share admitted to top three choice by subgroups (kernel density functions)

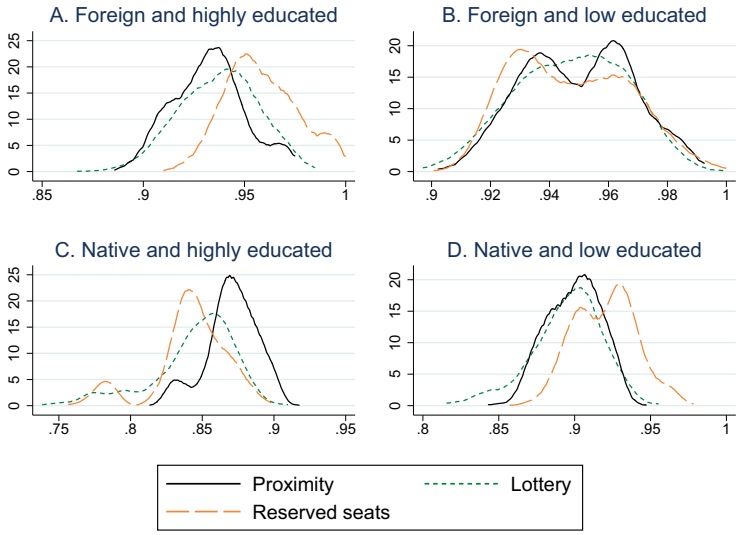
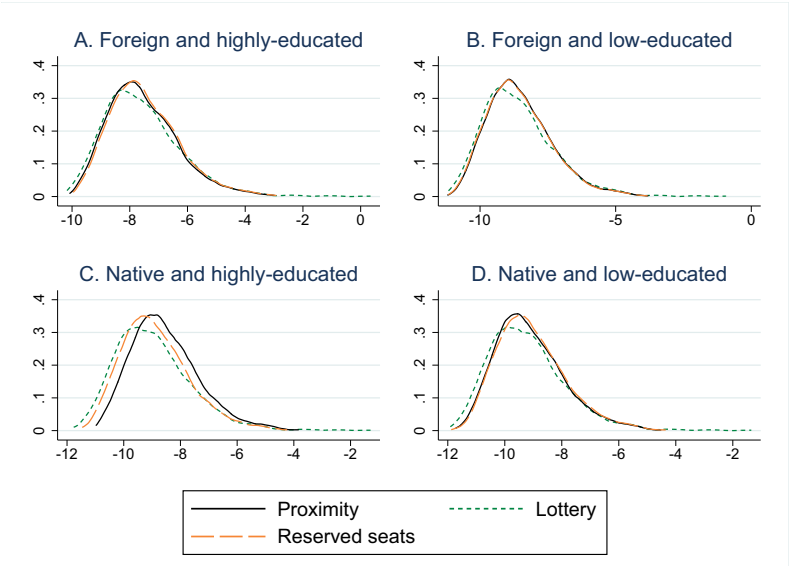


Figure 2.14: Utility by subgroup (kernel density functions)



Chapter 3

Debiasing the Gender Differences in Willingness to Compete

- The Effects of General Information on the Gender Gap and Efficiency*

*I would like to thank Nicklas Troedosn, Matilda Gullbransen Mattson, Arne Floryd and Johan Lundström, all teachers at Nacka high school, without whom this paper would not have been possible. I would also like to thank Anna Dreber Almenberg, Siri Isaksson, Karin Hederos, Anna Sandberg, Emma von Essen, Astri Muren and seminar participants at Stockholm University for valuable feedback and comments. Further I should thank Henrik Wallin who was with me when this idea was born and when the first round of experiments was conducted. I would also like to thank Mattias Hallberg who helped me conduct the second round of experiments.

3.1 Introduction

Over the last decade economists have recognized that gender differences in willingness to compete could be a contributing factor to the previously unexplained gender differences in, among other things, labor market outcomes. One of the first papers to address the question of whether women and men differ in their willingness to compete was Niederle and Vesterlund (2007). Niederle and Vesterlund (2007) performed a laboratory experiment where participants were asked to choose between a non-competitive piece-rate compensation scheme and a competitive tournament-based compensation scheme. They found that men were significantly more prone to select the competitive scheme, compared to women. Further, they documented that women, on average, competed too little to maximize expected payoff while men, on average, competed too much. In this sense, these gender differences in competitiveness are accompanied by an inefficiency. A large number of studies using different designs, in different settings, countries and cultures have confirmed these findings.[2]

[2]For example, Gneezy et al. (2009), Vandegrift and Abdullah (2009), Healy and Pate (2011) Dohmen and Falk (2011) Balafoutas and Sutter (2012) Balafoutas et al. (2012), Müller and Schwieren (2012), Dargnies (2012), Kamas and Preston (2012), Price (2012), Booth and Nolen (2012), Cárdenas et al. (2012), Kamas and Preston (2012), Mayr et al. (2012), Shurchkov (2012), Cadsby et al. (2013), Niederle et al. (2013), Gupta et al. (2013), (Andersen et al., 2013), Buser et al. (2014), Dreber et al. (2014), Lee et al. (2014), Sutter and Glatzle-Rutzler (2014), Khachatryan et al. (2015) and Buser et al. (2017a). For a detailed review of this literature see Niederle (2014). In most studies the gender differences in competition decreased but persisted when controlling for factors such as risk preferences and confidence. This has been interpreted to mean that the differences in the propensity to select into a competitive setting are, in large part, driven by gender differences in competitiveness as such and not by other underlying factors. This finding, that gender differences in competitiveness is its own phenomena, has been questioned in recent work by Gillen et al. (2015). Using new techniques and a unique data set they demonstrate that the gender differences in competitiveness could be due to measurement error in the experimental proxies, used to

Whether results from these laboratory experiments translate into real-world outcomes or not, has due to data limitations been notoriously hard to investigate. There is however a growing literature that has found ways to overcome these challenges. Most of these studies find that gender differences in the propensity to compete actually explain and/or predict gender differences in a variety of real-world outcomes, such as education choices, earnings, job search behavior, entrepreneurship and interest in political office.[3]

In her literature review Niederle (2014) discusses what the next steps in this research agenda should be. She writes:

> *A [...] question is whether these gender differences are indeed "true" differences in preferences, or whether they rather represent biases of women (or men), and whether awareness of those gender differences may therefore act as a way to "debias" the decisions of women and men.*[4]
> — Niederle (2014), page 8-9

measure risk preferences and confidence.

[3]Buser et al. (2014) study choice of secondary school track in the Netherlands and show that more than 20% of the gender gap can be explained by differences in competitiveness. Reuben et al. (2017), find that among NYU students, gender differences in overconfidence and competitiveness explains 18% of the gender gap in earnings expectations. Flory et al. (2015) conduct a large-scale field experiment where they randomize 9 000 job-seekers into different compensation regimes and find that women disproportionately shy away from competitive work settings. Berge et al. (2015) combine lab evidence on competitiveness with field evidence on investment, employment, profit, and sales. They find that competitiveness in the lab is positively associated with competitive choices in the field (investment and employment) and suggestive evidence of a positive link to successful field outcomes (profit and sales). Buser et al. (2017b) find that willingness to compete predicts career choices along the whole ability distribution and that the predictions differ for boys and girls. Preece and Stoddard (2015) conduct a field experiment and find that priming individuals to consider the competitive nature of politics has a strong negative effect on women's interest in political office, but not on men's.

[4]For a review on the literature on debiasing, see Soll et al. (2015).

This is the question I explore in this paper. I investigate if making people aware of the gender differences in the willingness to compete and the accompanying inefficacies can work to reduce said differences and inefficacies. I conducted a field experiment in twelve social science classes at a high school in Stockholm, where the participants had to solve simple math problems. The participants could choose if they wanted to be compensated by a competitive or a non-competitive compensation scheme. Six of the twelve classes were randomly selected as a treatment group. The participants in the treatment group were informed that (i) there are gender differences in the willingness to compete, with men, on average, competing more than women and (ii) that men, on average, compete too much given that they want to maximize their pay-off, while women on average, do not compete enough.

In the control group, the general findings from previous literature, are reproduced with men being about 30 percentage points more likely to compete than women, and men as a group, competing more than they should (assuming that they want to maximize expected pay-off) while the reverse is true for women. In the treatment group, the gender differences in competition are not only eliminated, but reversed, with women about 30 percentage points more likely to compete than men. Reversing the gender gap might not seem like a desirable outcome but in this setting, women are, on average, higher performing than men and therefore a higher fraction of women should compete (again assuming that they want to maximize expected pay-off). So, the treatment reduces inefficiency (measured as the proportion of incorrect choices to select or not select into competition) with 12 percentage points (corresponding to about 40%). If we take the compensation structure into account, inefficiency is reduced by almost 55%. This is because the treatment reduces costly mistakes more than less costly mistakes. The

effects on inefficiency differ between groups with low-performing men and high-performing women being the winners and low-performing women the losers. High-performing men are unaffected. I also find that the treatment makes women, especially high-performing women, more confident in their ability. This points to one of the potential mechanisms behind the treatment effects. Among men I do not find any effects on either risk preferences or confidence.

The results seem to imply that the gender differences in the willingness to compete are not due to innate and unchangeable preferences, but rather something that depends on the information an individual has available and the beliefs of that individual. The significant gains in efficiency of the treatment also suggest that information interventions of this type does not only have a positive effect on equality but also on efficiency. The fact that the gains are not distributed equally among groups does however create a trade-off that has to be taken into account when designing these information interventions. Possibly, the trade-off can be reduced by making the information more specific, for example by stating that it is high-performing women that seem to shy away from competition and not women in general.

This general conclusion is in line with a number of previous studies that have shown that the gender difference in the tendency to compete can be reduced. Cadsby et al. (2013) use a questionnaire in the beginning of the experiment to prime subjects concerning either gender/family or professional issues. They find that priming for professionalism significantly reduces the gender gap in tournament entry and that professional priming positively impacts women with respect to both their beliefs as well as their preference for tournament incentives. Wozniak et al. (2014) provide participants with information on their relative performance and show that this eliminates the gender gap and moves high-performing women towards more competitive

forms of compensation and low-ability men towards less competitive forms. In total, the average loss due to non-expected pay-off maximizing choices is reduced with about 25 %. Erttac and Szentes (2011) conduct a similar experiment and confirms the results. Niederle et al. (2013) pursue another avenue by exploring the effects of introducing affirmative action in form of a soft quota. They find that women are more prone to enter a tournament with a soft quota compared to one without, while the opposite is true for men. They do not find any effects on the performance of those choosing to compete implying that there are no efficiency losses. Balafoutas and Sutter (2012) conduct a similar experiment with similar results. They also consider an alternative affirmative action devise (preferential treatment where women's scores are increased by a point or two) and find that this also reduces the gender gap in tournament entry. Similar findings have been found in studies looking at single-sex experiments (Niederle, 2014). Compared to these interventions, the treatment in this paper seems to be effective in terms of increasing expected payoff. The intervention proposed in this paper is also easy to administrate and non-intrusive, compared to most other interventions.

The rest of the paper is organized as follows. Section 3.2 describes the experiment in detail, Section 3.3 presents the empirical strategy and the data, Section 3.4 presents the results and Section 3.5 concludes.

3.2 The Experiment

The experiment was conducted in classrooms at Nacka high school. The classrooms were set up in an exam setting preventing the participants from looking at each other's answers. The teachers were present during the entire experiment and the students not participating in the experiment were registered as absent. In total each

experiment, including instructions, took around 30 minutes.

The main part of the experiment was made up of three rounds of five minutes. In each of these rounds the participants were asked to solve as many simple math problems (adding five two-digit numbers) as they could during the allotted time. The problems were given to the participants on a worksheet and presented in the following form:

$$51 + 76 + 11 + 39 + 61 =$$

After each round the participants where given the answers to the questions allowing them to calculate their score. They where not given any information about the other participants' performances or their own relative performance.

What differed between the three rounds was the compensation scheme. In the first round the participants were rewarded with 5 SEK per correct answer (piece-rate compensation).[5] In the second round the participants were placed in a random group of four unknown peers.[6] If a participant had the highest number of correct answers in their group of four they were rewarded with 20 SEK per correct answer and if not, they received nothing (tournament-based compensation). In the third round, the participants could choose which of these two compensation schemes they would prefer to be compensated by. This choice was made just before the third round and will be used as an indicator for the participants' decision whether to compete or not.

After these three rounds there where two additional rounds where

[5]USD/SEK ≈ 8.33

[6]The participants were told that the peers were from the same population as them (high school students in the social science track) but where not given any additional information. The experiment was conducted in a trial group before the first actual experiment and the scores from this group were used to construct the peer groups. The performance distribution in this group more or less mirrors the performance distribution among the participants in the experiment.

confidence and risk preferences where elicited. In order to elicit confidence, the participants were asked to guess in what quartile they performed in the second round. A correct guess would yield 100 SEK.[7] To elicit risk preferences the participants were asked if they preferred a guarantee of 10/30/50/70/90 SEK or a 50% of winning 100 SEK.[8]

After these five rounds, the participants were asked to state their legal sex (as written in their passport) and thereafter two of the participants in each class were randomly selected to be compensated as described above.[9]

The main experiment was conducted in twelve social science classes at Nacka High School. Six classes were sampled in 2011 and six in 2014. In each year, three of the six classes sampled were randomly selected to be treated. The remaining six classes served as a control group. Before the third round of the experiment, the treated classes had the following information read to them:

Research has found that men are more prone to compete

[7]Again the performance in the trial group served to create the performance distribution they were compared to.

[8]No participants chose 50% chance of winning 100 SEK over a guarantee of 90 SEK. With some abuse of the concept I will refer to those always choosing the safe option as very risk-averse, those choosing the safe option in all but one case as somewhat risk-averse, those choosing the safe option in all but two cases as risk-neutral, those choosing the safe option in two cases as somewhat risk-seeking and those only choosing the safe option once as very risk-seeking. Note that this set-up allowed the participants to state inconsistent risk preferences, for example preferring a guarantee of 50 SEK to a 50% chance of wining a 100 SEK but, at the same time, not preferring a guarantee of 70 SEK to a 50% chance of 100 SEK. This was however not observed in the data.

[9]This was due to the, initially, very restricted budget of the project. The fact that not everyone was going to be compensated was conveyed to the participants from the start. The fact that only two participants were chosen for compensation makes the experiment low-yield for the participants but during the sessions, in general, the participants seemed to be focused on the task at hand.

than women. Further, men, on average, tend to compete too much given that they want to maximize their expected pay-off while women, on average, do not compete as much as they should.

3.3 Empirical Strategy

Treatment effects will be estimated by the following equation:

$$Y_i = \alpha + \beta T_i + \epsilon_i \tag{3.1}$$

where T_i is an indicator if class i was treated or not, Y_i is the share of students in class (or subgroup in class) i that selected into competition or that made a correct choice to compete and ϵ_i is an error term. β is the coefficient of interest. The regressions will be weighted with the relevant frequencies (the number of students in the class/subgroup in the class). Doing this gives β an intuitive interpretation: the treatment effect on the probability that an individual competes/makes the correct choice.[10]

Given randomization of treatment, strict exogeniety ($E[\epsilon|X] = 0$) should hold and hence β will be an unbiased estimator of the treatment effect. The small sample ($N = 12$) does however imply that we cannot rely on the asymptotic properties that underlie most conventional econometric methods for drawing inference. Inference regarding the treatment effects will therefore be based on randomization inference. In a setting where treatment is random, randomization inference provides an exact test of sharp hypotheses no matter the sample size (Young, 2017).

The motivating idea behind randomization inference is that, given

[10]The same interpretation as a linear probability model using the individual students in the classes as observations.

the sample of participants (in this case classes), the only stochastic element is the allocation of treatment. A null-hypothesis of no treatment effect on any of the participants ($Y_i(t_i) = Y_i(0) \quad \forall i$) can then be tested by calculating all possible realizations of a test statistic and rejecting the null if the observed realization, of the test statistic in the actual experiment, is extreme enough.

Randomization inference is described in detail in Young (2017). I will however give a simplified example relevant to this particular setting. In the data I observe that in all the six non-treated classes, men are more prone to compete than women while in all the treated classes, the reverse is true. The data generating process was such that in 2011 and 2014 respectively, three out of six classes were randomly chosen to be treated. We can calculate the probability that, given that there are no treatment effects on any of the classes, we happened to randomly select the six classes where women are more prone to compete than men, as the treatment group. This probability is given by:

$$(3/6 * 2/5 * 1/4)^2 = 1/400 = 0.0025$$

The probability that this pattern arose by chance is 0.25% and therefore we can comfortably reject that the null hypothesis (that there are no treatment effects on any of the treated observations) is true.

Reality is however more complex than in the stylized example above and even with a small sample size the number of possible perturbations of treatment increases rapidly. Calculating the outcome for each possible perturbation is therefore tedious. Hence, I will rely on simulations. I will resample my treatment variable 1000 times for each outcome of interest and subgroup.[11] For each perturbation

[11] 200 repetitions have been shown to be enough for the p-value to stabilize (Young, 2017).

Regression 3.1 is estimated and the corresponding β is documented. The fraction of perturbations that yielded a more extreme value of β than the one observed in the actual experiment ($\beta_{obs} > \beta_p$ when $\beta_{obs} > 0$ and $\beta_{obs} < \beta_p$ when $\beta_{obs} < 0$) is then computed, this fraction is what Young (2017) refers to as the randomized p.[12] This is implemented in Stata using the ritest command described in HeB (2017). As three classes where sampled into the treatment group in 2011 and three in 2014, I will resample using the year of treatment as a stratum.

A total of 12 classes were sampled into either the control group or the treatment group. Summary statistics for these groups can be found in Table 3.1. At the day of the experiment the average number of students in the classrooms were 22.33. The average share of women was 0.56. The average score for men was 5.30 in round one and 6.17 in round two. The corresponding numbers for women were 6.46 and 6.48. Overall women outperformed men which, in math, is unusual. Sweden is however an outlier when it comes to the gender gap in math performance.[13] We also observe the well-documented pattern that men improve under competition while women do not.

We can also see that the treatment group and the control group seem to be very similar in terms of observable characteristics. This should not be surprising as allocation to treatment was randomized. To help convince the reader that the randomization was not compromised, Table 3.2 reports the results of a regression with the treatment indicator as the dependent variable and all the variables in Table 3.1 as dependent variables. None of the estimates are individually close

[12]Note that it is possible for the β in a given perturbation to be the same as in the observed data. This usually happens when the treatment group in the perturbation is the same as in the actual sample or when both the outcome and the treatment variable are binary variables. As I use strict inequalities, these cases will be counted towards acceptance so as not to over-reject.

[13]See, for example, PISA 2015

Table 3.1: Summary Statistics

	Total (N=12)		Control (N=6)		Treatment (N=6)	
	Mean	SD	Mean	SD	Mean	SD
Number of students	22.33	(3.06)	22.83	(3.45)	21.83	(2.78)
Share of women	0.56	(0.16)	0.57	(0.16)	0.54	(0.18)
Grade	2.00	(0.85)	2.16	(0.98)	1.83	(0.75)
Performance round one (men)	5.30	(0.98)	5.51	(0.92)	5.09	(1.08)
Performance round one (women)	6.46	(0.96)	6.52	(1.21)	6.39	(0.77)
Performance round two (men)	6.17	(0.76)	6.23	(0.59)	6.10	(0.96)
Performance round two (women)	6.48	(0.87)	6.49	(1.04)	6.47	(0.76)

Note: Standard deviations in parenthesis.

to being statistically significant and an F-test shows that they are not jointly significant.

Table 3.2: Test of Randomization

	Treated
Number of students	-0.050
	(0.482)
Share of women	0.165
	(0.936)
Grade	0.120
	(0.598)
Performance round one (men)	-0.196
	(0.613)
Performance round one (women)	-0.189
	(0.529)
Performance round two (men)	0.153
	(0.744)
Performance round two (women)	0.063
	(0.825)
Observations	12
P-value from F-test of joint significance	0.996

*Note: Randomization inference p-values in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

3.4 Results

Column 1 in Table 3.3 reports the tendency to compete in the control group for men and women respectively. About 56% of men compete while only 25% of women do. This large and significant gap in the tendency to select into competition is in line with previous findings.

Column 2 and 3 report the average confidence of the participants. 56% of men and 47% of women correctly manage to guess their performance quartile. Among men, overconfidence is about as common as underconfidence (20/24%). Among women it is much more common to be underconfident (40%) than overconfident (13%). The participants in the control group are somewhat less confident than in most previous literature but the general pattern that men are more confident than women is confimed. The last two columns report the risk preferences. Again, men are more risk-taking and less risk-averse than women.

If we assume that performance in round two is a good predictor for performance in round three[14] it is possible to calculate the expected pay-off of competing and not competing for each participant. These expected pay-offs can be used to define if a participant makes a "correct" decision to compete or not. Let a correct decision be defined as competing when the expected pay-off of competing is higher than the expected pay-off from not competing or not competing when the reverse is true. The performance cut-off for when the expected pay-off is higher from competing than not competing is eight correct answers in round two.[15] Using the expected pay-off, I can also calculate the actual expected loss from these incorrect decisions.

Table 3.4 reports these statistics for the control group. Panel A presents the results for the entire sample. 71% of the participants make a correct decision. About half of the incorrect decisions are those where participants compete when they should not, and half are when participants do not compete when they should. We can also see that the average loss, due to incorrect decisions is 7.71 SEK and that this is mainly driven by those deciding not to compete when

[14]The correlation is 0.83.

[15]This is the case in the whole sample, in the control group and the treatment group separately as well as in the trial group.

Table 3.3: Confidence and Risk Preferences in the Control Group

	Compete	Overconfident	Underconfident	Risk-seeking	Risk-averse
Men (N=59)	0.56	0.24	0.20	0.12	0.64
Women (N=78)	0.25	0.13	0.40	0.04	0.81
Diff (Men-Women)	0.31	0.11	-0.19	0.08	-0.16

Note: A participant is defined as being overconfident if they guess that they are in higher performance quartile than they actually are and underconfident if they guess that they are in lower performance quartile than they actually are. Those always choosing the safe option and those choosing the safe option in all but one case in round five are defined as risk-averse while those choosing the safe option in just one or two cases are defined as risk-seeking.

they should. This is due to the structure of the compensation scheme, with the rewards per correct answer being four times larger under the tournament-based compensation scheme, compared to the piecerate compensation scheme.

Panel B and Panel C report the same statistics for the subset of men and women respectively. 68% of men and 73% of women make correct decisions. The reason for the incorrect decisions varies greatly between men and women. Among men 85% of incorrect decisions are due to them competing when they should not, while among women 81% of incorrect decisions are due to them not competing when they should. This is also reflected in the average loss. In general, this is the pattern found in previous research, men, as a group, compete too much while women shy away from competition.

3.4.1 Treatment Effects

Table 3.5 reports the results from Regression 3.1 with the share of participants in each class choosing to compete as the dependent variable. We can see that the treatment has no aggregate effect. It does however have a large effect when we look at men and women separately. Men become 30 percentage points less likely to compete while women become 29 percentage points more likely to compete. The initial gender difference of 31 percentage points is almost reversed.

Table 3.6 reports the results from Regression 3.1, with the share of participants in each class making a correct choice as the dependent variable. In the first column we can see that in total the share of correct choices increases with 11.6 percentage points. This corresponds to a reduction of incorrect choices of around 40%. The second column reports the results from running the same regression, but with the expected loss as the dependent variable. The expected loss is reduced by 4.2 SEK or about 55%. The third column uses the performance in round three instead of the performance in round two to

Table 3.4: Performance, Correct Choices and Expected Loss

A: All participants (N=137)

	Should compete	Should not compete	Total
Competes	0.23	0.15	0.38
Does not compete	0.15	0.48	0.63
Total	0.38	0.63	

	From competing	From not competing	Total
Average loss	2.68	5.03	7.71

B: Men (N=59)

	Should compete	Should not compete	Total
Competes	0.29	0.27	0.56
Does not compete	0.05	0.39	0.44
Total	0.34	0.66	

	From competing	From not competing	Total
Average loss	4.79	1.63	6.42

C: Women (N=78)

	Should compete	Should not compete	Total
Competes	0.19	0.05	0.24
Does not compete	0.22	0.54	0.76
Total	0.41	0.59	

	From competing	From not competing	Total
Average loss	1.08	7.6	8.68

Note: Participants that should compete are those with a performance of eight or more in round two. Average loss is calculated based on the pay-off a participant is expected to receive under the different compensation schemes in round three given their performance in round two.

Table 3.5: Effect on Tendency to Compete

	All	Men	Women
Treatment	0.033	-0.301***	0.291***
	(0.363)	(0.009)	(0.004)
Constant (baseline)	0.400	0.559	0.244
Observations	12	12	12

*Note: Randomization inference p-values in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

evaluate if a choice was correct or not. In doing this we are moving from expected loss to realized loss. The results from column one are unchanged. In the fourth column all participants close to the cut-off are removed from the sample. This means that we exclude the cases where the expected pay-off does not differ much between competing and not competing. Again, the results from column one are more or less unchanged. The fifth column combines the restriction from column four and uses performance in round three to establish the cut-off. If anything, the effects here, are even larger. Hence, this considerable gain in efficiency seems to be robust to the definition of a correct decision.

Table 3.7 reports the effects on the share of correct choices and expected loss on subgroups. Panel A reports the results for the share of correct choices. We can see that there are both winners and loosers from the treatment. The big winners are low-performing men that increase their likelihood of making a correct choice by 47.5 percentage points and high-performing women that increase their probability of making a correct choice by 33.9 percent. These two groups are the ones that were responsible for the lion's share of incorrect choices in

Table 3.6: Effects on Correct Choices

	All	All	All	All	All
Treatment	0.116**	-4.236**	0.117**	0.107**	0.143***
	(0.032)	(0.015)	(0.014)	(0.028)	(0.004)
Constant (baseline)	0.708	7.7	0.715	0.747	0.748
Observations	12	12	12	12	12
Performance in round three	No	No	Yes	No	Yes
Without close cases	No	No	No	Yes	Yes

*Note: Randomization inference p-values in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

the control group. The main losers are low-performing women, whose likelihood of making a correct decision decreases by 21.7 percentage points. High-performing men seem to be unaffected. The opposite effects on high-performing and low-performing women mean that, in total, there are no effects on women.

In Panel B the effects on the expected loss are reported. We can see that when we take the actual structure of the compensation scheme into account the results differ somewhat. High-performing women are now the big winners with a reduction in expected loss of 18.3 SEK. This big reduction in expected loss is now more than enough to compensate for the increased expected loss among low-performing women of 3.62 SEK, meaning that women as a group benefit from the treatment. Again, high-performing men are unaffected.

Table 3.7: Effects on Correct Choices, Subgroups

	Men			Women		
	All	High-performing	Low-performing	All	High-performing	Low-performing
			A: Correct Choices			
Treatment	0.201**	-0.100	0.475***	0.050	0.339***	-0.217**
	(0.023)	(0.459)	(0.002)	(0.289)	(0.003)	(0.011)
Constant (baseline)	0.678	0.850	0.469	0.731	0.590	0.913
			B: Expected loss			
Treatment	-3.02	4.12	-5.95***	-5.15**	-18.28***	3.62***
	(0.143)	(0.333)	(0.008)	(0.025)	(0.003)	0.006
Constant (baseline)	6.41	4.81	7.23	8.67	18.50	1.84
Observations	12	12	12	12	12	12

*Note: Randomization inference p-values in parenthesis. Significance levels indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

3.4.2 Does Gender Matter?

The information received by the treatment group contains two different components. The first component can be thought of as a warning to the participants that one can be too competitive or too careful given that one wants to maximize the expected pay-off. The second component is that this is related to gender. The question we are concerned with in this project is the effects of information on gender differences as such. Therefore we want to isolate the effect of the first component from the effect of the two components combined. Another set of three experiments was therefore conducted in 2016 where the following information was read to the participants:

> *Research has found that some people tend to be too prone to compete, given that they want to maximize their expected pay-off while others, do not compete as much as they should.*

This group of three classes will be referred to as the warning group. The warning group was not randomly selected from the same sample as the other two groups but rather added on later. This means that the strategy used to estimate causal effects and draw inference used in this paper is not applicable. In fact, without strong distributional assumptions there is no credible way of drawing statistical inference on the differences between this group and the other two groups. I will therefore limit myself to just comparing means in the different groups.

Table 3.8 reports how the warning group differs from the control and treatment groups in terms of observable characteristics. In general there are no big differences but we can see that the classes in the warning group have a few more students in each class than the other groups. In general, the similarity between the groups is reassuring as

it indicates that the warning group is not sampled from a completely different population and can serve our purposes.

In Figure 3.1 we can see how the mean in the control group differs from the means in the treatment group and warning group respectively. We see that in almost all cases the warning group deviates from the control group in the opposite way from how the treatment group deviates from the control group. For example, in the treatment group, men are less prone to compete than in the control group, while in the warning group the reverse is true. The only case where this is not true is the total tendency to compete where the participants, when compared to the control group, are more prone to compete in both the warning and the treatment group. These differences are however small and as is shown in Table 3.5, the treatment effect in question is not statistically significant. Given this, it seems unlikely that it is the warning as such that is the underlying reason for the observed treatment effects. If anything, it indicates that the warning alone could have the opposite effect to the composite treatment (the treatment group). Given the limited number of observations and the problem with causal inference, this is however left for future research to investigate further.

3.4.3 Mechanisms

Gender differences in confidence and risk have been shown to explain, at least part of the gender differences in the tendency to select into competition. In round four and five of the experiment, proxies for risk preferences and confidence were elicited and so we can explore if the treatment seem have had any effect on either of these potential underlying factors.

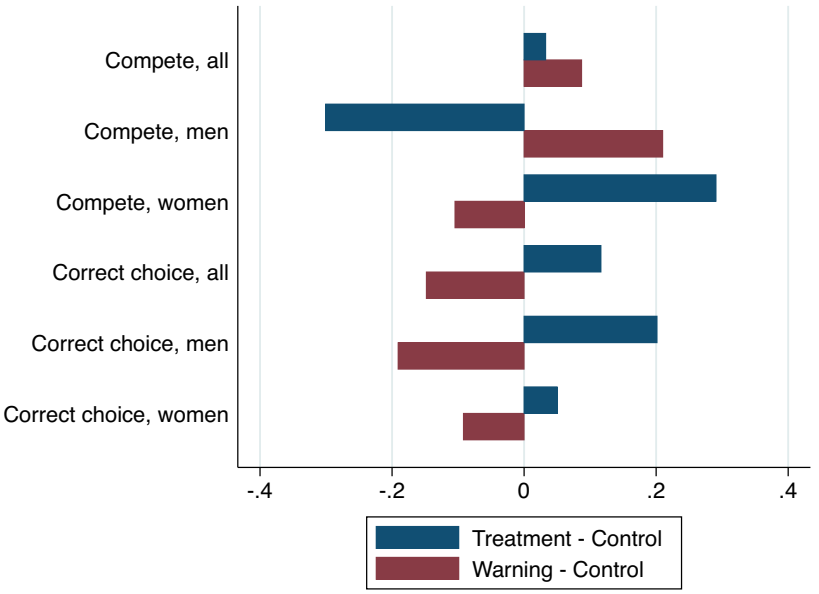
Figure 3.2 compares the cumulative distributions of the participants guessed performance quartile (with one being the lowest performing quartile and four being the highest) for men and women in

Table 3.8: Summary Statistics, Control Group, Treatment Group and Warning Group

	Control (N=6)		Treatment (N=6)		Warning (N=3)	
	Mean	SD	Mean	SD	Mean	SD
Number of students	22.83	(3.45)	21.83	(2.78)	25	(3.61)
Share of women	0.57	(0.16)	0.54	(0.18)	0.48	(0.21)
Grade	11.16	(0.98)	10.83	(0.75)	10.33	(0.58)
Performance round 1 (men)	5.51	(0.92)	5.09	(1.08)	5.09	(0.81)
Performance round 1 (women)	6.52	(1.21)	6.39	(0.77)	6.10	(0.96)
Performance round 2 (men)	6.23	(0.59)	6.10	(0.96)	6.40	(0.58)
Performance round 2 (women)	6.49	(1.04)	6.47	(0.76)	6.60	(0.29)

Note: Standard deviations in parenthesis.

Figure 3.1: Difference from Control Group, Treatment Group and Warning Group



Note: This figure shows the differences between the control group and the treatment group and the warning group respectively in various outcomes.

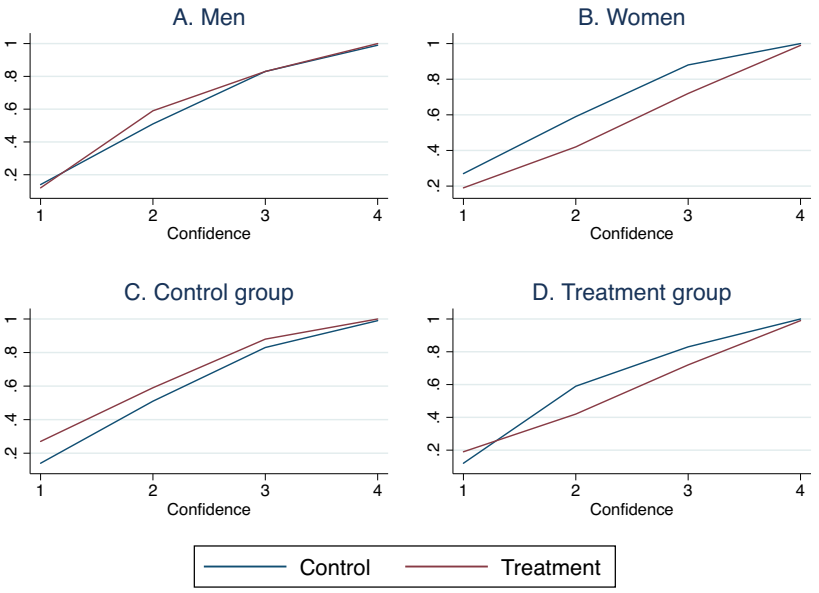
the treatment group and control group respectively. In Subfigure A, men in the control group are compared to men in the treatment. We see that the distributions do not differ considerably. The share of men that correctly guess their performance quartile is also similar between the groups (56% in the control group and 57% in the treatment group). In Subfigure B, this exercise is repeated for women. We see that the treatment gives women considerably more confidence (Randomized $P = 0.004$). The share of women correctly guessing their performance quartile also increases by 20 percentage points (from 47% in the control group to 67% in the treatment group, randomized $P = 0.007$). This effect is entirely driven by high-performing women from whom the share of correct guesses increases by 50 percentage points. All together, this means that women go from being less confident than men (as seen in Subfigure C) to being more confident than men (as seen in Subfigure D). Women also improve their ability to guess their performance rank and, in the treatment group, they outperform men in this respect.

Figure 3.3 shows similar cumulative distributions for risk preferences (with one being the most risk-averse and five being the most risk-loving). We can see that men have stronger preferences for risk compared to women and that the risk preferences of neither men nor women seem to be affected markedly by the treatment.

In general, this seems to imply that increased confidence among women is one explanation for the reduced gender gap in the tendency to select into competition. It does not however seem to explain everything. For men we saw a large decrease in the tendency to select into competition, but no effects on either risk preferences or confidence. Further, we saw that the treatment made men more likely to make a correct choice to compete or not, but we do not see any evidence of them becoming better at guessing their own performance quartile. This could mean that there are other mechanisms at play, but

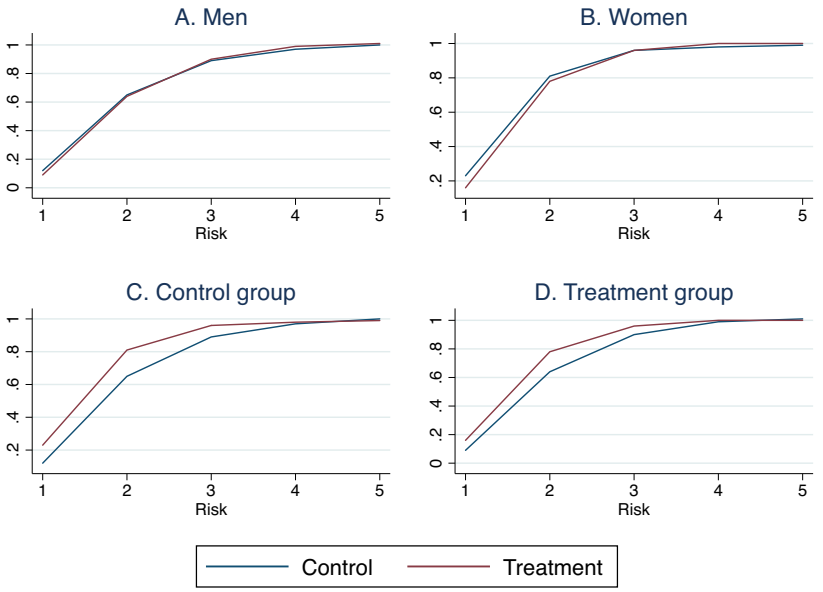
it could also be because the methods used to elicit risk preferences and confidence are too crude (as discussed by Gillen et al. (2015)).

Figure 3.2: Cumulative Distributions of Confidence



Note: This figure shows the cumulative distribution of the participants guessed performance quartile for men and women in the treatment group and control group respectively. One represents the lowest performing quartile and four the highest perfuming quartile.

Figure 3.3: Cumulative Distributions of Risk Preferences



Note: This figure shows the cumulative distribution of risk-preferences for men and women in the treatment group and control group respectively. Risk-preferences are measured on a five point scale with one being the most risk-averse and five being the most risk loving.

3.5 Concluding Remarks

This paper has explored if information on gender differences willingness to compete and the accompanying inefficiencies can reduce said differences and inefficiencies. An experiment where the participants could choose whether to select into a competitive setting or not was conducted at a Swedish high school. Out of twelve classes, six were randomly selected into a treatment group that were informed about the gender differences in the willingness to compete and the related inefficiencies before making their choice. In the control group, men were much more likely to select into competition, compared to women. There were also significant inefficiencies from low-performing men choosing to compete and high-performing women choosing not to. In the treatment group this gender gap was reversed and there were significant gains to efficiency. There are however winners and losers of the intervention. The big winners are low-performing men and high-performing women and the big losers are low-performing women.

Compared to previous explored ways of reducing the gender gap in the willingness to compete, this intervention seems to be efficient. It is also non-intrusive and easy to administrate. Given this, it seems like a policy that deserves further exploration. There are four main questions that need to be answered. The first is if the effect of the treatment is dependent on the level of awareness of gender norms. A social science class at a Swedish high school is probably a setting where this awareness, on average, is very high. An example of this is that I was warned by the teachers about asking the students about their gender without making clear that I was asking for their legal gender, as it might cause a debate and disrupt the experiment. Secondly, the experiment should be replicated in settings where women are not outperforming men as the efficiency gains documented in this

setting might be contingent on this being the case. Thirdly, as with all experiments, it is important to show that these effects translate to real-world outcomes. Related to this third question, is the fourth question of the longevity of the effects. In this case, treatment was administrated just before the choice to compete or not was made. In most potential real-world applications, it could prove hard to provide this sort of information in direct conjunction with the decision to select into a competitive environment or not. What would the effects have been if treatment had been administrated a week/month/year beforehand? And can repeated information interventions increase the longevity of the effect?

Bibliography

Andersen, S., Ertac, S., Gneezy, U., List, J. A., and Maximiano, S. (2013). Gender, Competitiveness, and Socialization at a Young Age: Evidence From a Matrilineal and a Patriarchal Society. *Review of Economics and Statistics*, 95(4):1438–1443.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional Preferences and Competitive Behavior. *Journal of Economic Behavior & Organization*, 83(1):125–135.

Balafoutas, L. and Sutter, M. (2012). Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory. *Science*, 335(579):579–582.

Berge, L. I. O., Bjorvatn, K., Pires, A. J. G., and Tungoddena, B. (2015). Competitive in the Lab, Successful in the Field? *Journal of Economic Behavior & Organization*, 118:303–317.

Booth, A. and Nolen, P. (2012). Choosing to Compete: How Different are Girls and Boys? *Journal of Economic Behavior & Organization*, 81(2):542–555.

Buser, T., Dreber, A., and Mollerstrom, J. (2017a). The Impact of Stress on Tournament Entry. *Experimental Economics*, 20(2):506–530.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, Competitiveness and Career Choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Buser, T., Peter, N., and Wolter, S. C. (2017b). Gender, Willingness to Compete and Career Choices along the Whole Ability Distribution. *IZA Discussion Papers No. 10976*, 10976. IZA - Institute of Labor Economics.

Cadsby, C. B., Servátka, M., and Song, F. (2013). How Competitive are Female Professionals? A Tale of Identity Conflict. *Journal of Economic Behavior & Organization*, 92:284–303.

Cárdenas, J.-C., Dreber, A., von Essen, E., and Ranehill, E. (2012). Gender Differences in Competitiveness and Risk Taking: Comparing Children in Colombia and Sweden. *Journal of Economic Behavior & Organization*, 83(1):11–23.

Dargnies, M.-P. (2012). Men Too Sometimes Shy Away from Competition: The Case of Team Competition. *Management Science*, 58(11):1982–2000.

Dohmen, T. and Falk, A. (2011). Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender. *American Economic Review*, 101(2):556–590.

Dreber, A., von Essen, E., and Ranehill, E. (2014). Gender and Competition in Adolescence: Task Matters. *Experimental Economics*, 17(1):154–172.

Erttac, S. and Szentes, B. (2011). The effect of performance feedback on gender differences in competitiveness: Experimental evidence. *Koc University-TUSIAD Economic Research Forum Working Papers No. 1104*. Koc University-TUSIAD Economic Research Forum.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions . *The Review of Economic Studies*, 82(1):122–155.

Gillen, B., Snowberg, E., and Yariv, L. (2015). Experimenting with measurement error: Techniques with applications to the caltech

cohort study. *NBER Working Paper No. 21517*. National Bureau of Economic Research, Inc.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society. *Econometrica*, 77(5):1637–1664.

Gupta, N. D., Poulsenm, A., and Villeval, M.-C. (2013). Gender Matching and Competitiveness: Experimental Evidence. *Economic Inquiry*, 51(1):816–835.

Healy, A. and Pate, J. (2011). Can Teams Help to Close the Gender Competition Gap? *The Economic Journal*, 121(555):1192–1204.

HeB, S. (2017). Randomization Inference with Stata: A Guide and Software. *The Stata Journal*, 17(3):630–651.

Kamas, L. and Preston, A. (2012). Gender and Social Preferences in the US: An Experimental Study. *Feminist Economics*, 18(1):135–160.

Khachatryan, K., Dreber, A., von Essen, E., and Ranehill, E. (2015). Gender and Preferences at a Young Age: Evidence from Armenia. *Journal of Economic Behavior & Organization*, 118:318–330.

Lee, S., Niederle, M., and Kang, N. (2014). Do Single-Sex Schools make Girls more Competitive? *Economics Letters*, 124(3):474–477.

Mayr, U., Wozniak, D., Davidson, C., Kuhns, C., and Harbaugh, W. T. (2012). Competitiveness Across the Life Span: The Feisty Fifties. *Psychology and Aging*, 27(2):278–285.

Müller, J. and Schwieren, C. (2012). Can Personality Explain what is Underlying Women,s Unwillingness to Compete? *Journal of Economic Psychology*, 33(3):448–460.

Niederle, M. (2014). Gender. *NBER Working Paper No. 20788*. National Bureau of Economic Research, Inc.

Niederle, M., Segal, C., and Vesterlund, L. (2013). How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness. *Management Science*, 59(1):1–16.

Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away from Competition? Do Men Compete too Much? *Quarterly Journal of Economics*, 122(3):1067–1101.

Preece, J. and Stoddard, O. (2015). Why Women don,t Run: Experimental Evidence on Gender Differences in Political Competition Aversion. *Journal of Economic Behavior & Organization*, 117:296–308.

Price, C. R. (2012). Gender, Competition, and Managerial Decisions. *Management Science*, 58(1):114–122.

Reuben, E., Wiswall, M., and Zafar, B. (2017). Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender. *The Economic Journal*, 127(604):2153–2186.

Shurchkov, O. (2012). Under Pressure: Gender Differences in Output Quality and Quantity under Competition and Time Constraints. *Journal of the European Economic Association*, 10(5):1189–1213.

Soll, J. B., Milkman, K. L., and Payne, J. W. (2015). A user's guide to debiasing. In Keren, G. and Wu, G., editors, *Wiley-Blackwell Handbook of Judgment and Decision Making*. John Wiley and Sons, Ltd.

Sutter, M. and Glatzle-Rutzler, D. (2014). Management Science. *Gender Differences in the Willingness to Compete Emerge Early in Life and Persist*, 61(10):2339–2354.

Vandegrift, D. and Abdullah, Y. (2009). Men, Women, and Competition: an Experimental Test of Behavior. *Journal of Economic Behavior & Organization*, 72(1):554–570.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1):161–198.

Young, A. (2017). Channelling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. url: http://personal.lse.ac.uk/YoungA/ChannellingFisher.pdf (accessed 1 April 2018).

Chapter 4

The Housing Wealth Effect: Quasi-Experimental Evidence*

*This chapter was was co-authored by Björn Tyrefors Roine Vestman. We thank Matz Dahlberg, Peter Fredriksson, Soren Leth-Petersen, Kurt Mitman, Morten Ravn, Kathrin Schlafmann and seminar participants at Stockholm University, the Greater Stockholm Macro Group, the Research Institute of Industrial Economics (IFN), the Swedish House of Finance, and Sveriges Riksbank. We are grateful for generous funding from Handelsbanken's Research Foundation and from the Royal Swedish Academy of Sciences. All data used in this research have passed ethical vetting at the Stockholm ethical review board and have been approved by Statistics Sweden.

> *[...] there are good reasons to be skeptical of results based on macroeconomic or regional data (including our own). Foremost among these is the previously-acknowledged point that movements in asset prices are not exogenous fluctuations; they should be affected by many of the same factors that affect consumption decisions, most notably overall macroeconomic prospects. [...] The perfect experiment observed in the perfect microeconomic dataset is not available.*
> — Carroll et al. (2010)

4.1 Introduction

How much households consume out of housing wealth has been of major interest to economists for a long time. The reason is obvious: most households have a large share of their total wealth tied in housing, implying that fluctuations in house prices or housing wealth, even if the consumption elasticity is modest, has the potential to fuel or dampen aggregate household spending (see, e.g., Iacoviello, 2011).

Early studies that estimated the elasticity to consume out of housing wealth exploited time-series variation in macroeconomic series. Identification of the elasticity in such data relies on adding the correct control variables in order to avoid that the effect of macroeconomic conditions on consumption is confounded with the effect of house prices.[2] Reliance on micro data rather than data from the national accounts does not provide a satisfactory solution.[3]

[2]See e.g. Carroll (2004), Case et al. (2005), Calomiris et al. (2012) and Carroll et al. (2011) for a study with structural identification.

[3]For examples of well-crafted micro data studies, see Campbell and Cocco (2007), Browning et al. (2013), and Disney et al. (2010), and Attanasio et al. (2011).

Arguably, the best studies to date were produced as a response to the financial crisis of 2008-2009 and the booming housing market and lax lending standards that led to a subsequent housing bust. Mian et al. (2013) and Mian and Sufi (2014) use zip and county level data to quantify the relationship between housing wealth and consumption. Specifically, they use instrumental variables to isolate the component of the shock to house prices in a region (country or ZIP code area) that depend on geographic determinants and hence housing supply (Saiz, 2010). Mian et al. (2013) finds a short-run MPC out of housing wealth on cars equal to 2.3 cents per dollar loss of housing wealth and an MPC on total consumption amounting to 5-7 cents per dollar. Aladangady (2017) refines the instrumental variable by combining the housing supply elasticity with changes to long-term real interest rates, and use the U.S. Consumer Expenditure Survey on households from 1986 to 2008 to estimate an MPC of 4.7 cents per dollar among homeowners.[4]

This paper raises the standards of the literature in two ways. First, it meets even stricter requirements on establishing a casual link between a house price movement and its effect on household consumption. Our identification strategy relies on a quasi-experiment, in the form of the unanticipated continued operation of Stockholm's city airport. The airport, Bromma Airport, is surrounded by residential housing and there has been continuous debate over the negative externalities that the airport imposes on its surroundings, not the least in terms of noise. Due to its inappropriate geographic location, Bromma was meant to be closed in 2011.[5] However, in September 2007 the municipality of Stockholm unexpectedly offered a new con-

[4]For a critique of the basic uses of the supply elasticity, see Davidoff (2013, 2016).

[5]The main airport in the Stockholm region is Arlanda Airport and it is just 30 kilometers away from the city. This implies that the plan to shut down Bromma Airport was plausible.

tract spanning over 30 years to Bromma Airport, thus postponing the closing of the airport by 26 years. We use this unanticipated news to isolate a house price fall close to the airport. Thus, our identification strategy relies neither on long time series, nor on well-measured control variables, nor on time-invariant local geographic conditions. Instead, it relies on house price variation in the spatial, or geographic, domain within just four quarters (2007Q4 – 2008Q3) due to an unexpected political change. A short time window is important for a causal interpretation. We estimate a difference-in-difference regression on the universe of housing transactions in Stockholm to establish that house prices close to the airport fell by about 17% relative to prices of houses further away, following the unanticipated renewal of the contract. For a residential homeowner close to the airport this house price fall amounts to a relative loss in housing wealth equal to about SEK 469,000 ($\approx$ USD 59,000).[6] The first stage is strong and passes the standard test for parallel trend indicating that the instrument is as good as randomly assigned.

Second, we measure spending as well as loan-to-value ratios at the household level. We relate car consumption to the extension of the contract using a detailed household-level panel data set that, in addition to geographic information, includes sociodemographic information, complete information on households' balance sheets – including the loan-to-value ratio – as well as car transactions. Notably, all our data is at the household level and is matched from different administrative registers, implying that it is of excellent quality. Using car purchases as a proxy for household consumption, we find that homeowners close to the airport do not respond on the extensive margin (i.e. the number of car transactions close the airport relative to number of car transactions further away is unaffected). In addition, we find no evidence of a substitution effect to cheaper used cars or be-

[6]The exchange rate is about 8 SEK/USD.

tween new and used cars. However, homeowners close to the airport respond by reducing the value of new cars purchased by 7.2%. The effect is thus solely concentrated to the intensive margin among purchasers of new cars. Our interpretation of this finding is that of a substitution effect – households buy cheaper but not fewer cars. The intensive margin effect is robust and economically significant. We use a TSLS approach to esimate the elasticity and find that it ranges from 0.3-0.5 among households that purchase a new car. Converted to an MPC this equals 4.3 cents on every dollar. However, these estimates pertain only to purchasers of new cars. Therefore the aggregate effect among *all homeowners* is an order of magnitude smaller.

In addition, we observe balance sheets and (combined) loan-to-value ratios at the household level in 2006, prior to the renewal of the airport's contract. We use this information to document substantial cross-sectional heterogeneity in effects. We find a substantially greater effect among households with a loan-to-value (LTV) in the range of 0.6 to 0.8. This empasizes the role of credit and mortgage markets for how house price fluctuations affect the macroeconomy also in normal economic times.[7]

We emphasize four reasons to believe that our estimate is precise and credible. First, the difference-in-difference approach, as such, is a perfect marriage between diagnostics of parallel trends prior to the occasion of the quasi-experiment and asset pricing. Parallel trends in house price growth close and far away from the airport, prior to

[7]To the best of our knowledge, this is the first paper to document heterogeneity in response to house prices across households with different LTVs using household level registry-based data. Aladangady (2017) uses survey-based data. Often estimates rely on aggregated data. Aggregation to a higher geographical area, such as ZIP code, would be problematic if the housing wealth effect displays a non-linear relationship over the cross-sectional distribution of households' LTV ratios. Aggregation to a higher geographical area can sometimes also lead to other complications of the interpretation of estimates. See Adelino et al. (2015) and Adelino et al. (2016) for a discussion.

the renewal of the airport's contract imply that our identification strategy passes the test whether the contract renewal was in fact an unanticipated shock to house prices. Second, we detect spatial, or geographic, variation in the house price shock. Homeowners close to the airport are given larger doses of the negative shock than those further away and respond more in terms of car purchases. This speaks directly to the exclusion restriction assumption of an instrumental variables estimate, since in order for it to be invalid we need a time changing variable that is changing in September 2007 and is correlated to the distance of the elliptic noise contour and is affecting the consumption decisions of households. Any of these statements alone will not violate the exclusion restriction, only the combination. We think it is hard to find such variables, in particular when extending the model to be dose dependent and finally be estimated by continuous treatment. Third, we do not at all aggregate our data but are able to use disaggregated household data which allows us to directly speak to individual households' collateral constraints and loan-to-value ratios. Fourth, our design only relies on relative price movements. Economic conditions in Stockholm were good up until 2008Q3. The Stockholm housing market in general, was in a boom during the years of our study. However, any attempt to calibrate a macroeconomic model to our estimates must bear this in mind, i.e. that our elasticity is measured in a setting with a combination of a booming housing market and local negative house price shock may have different implications for tests of ambitious macroeconomic models such as Berger et al. (2017), Gorea and Midrigan (2017), and Kaplan et al. (2017). Nevertheless, we believe that our estimates are of great value to the literature since our estimate is independent of general equilibrium effects, such as labor income effects or behavioral effects.[8]

[8]See footnote 12 in Berger et al. (2017) for a discussion.

The rest of the paper is structured as follows. Section 4.2 reviews the institutional setting and the quasi-experiment, Section 4.3 describes the data and Section 4.4 discusses the empirical strategy. Section 4.5 reports the results, and Section 4.6 concludes.

4.2 Institutional Setting and the Quasi-Experiment

This section provides an overview of Bromma Airport's history and explains the quasi-experiment.

4.2.1 History and Political Governance

Bromma Airport is the Stockholm's city airport and the third lagest in Sweden. The airport has one runway and is located close to the city in an area that is otherwise dominated by single-family housing. In the years between 2006 and 2015 it had about 60,000 takeoffs and landings per year. In the early years, Bromma was Sweden's largest airport but after the establishment of the major international airport of Arlanda in 1959, Bromma airport saw a sharp decrease in traffic. By 1983 the traffic mainly consisted of private aviation but 1992 the conservative-liberal national government opened up commercial airfare to competition and Bromma airport experienced an increase in traffic again. During the late 1990s and early 2000s, however, there was a general perception that Bromma airport would be closed – by latest 2011 when the operating contract was set to expire.[9]

[9]There was a series of reports planning for the shutdown. In 1989 the municipality of Stockholm presented a major report proposing closing the airport by 1996 and using the centrally located area for housing (Stockholm, 1989). In 1994 the national government put together a commission tasked with the objective to figure out how fast Bromma airport could be phased out (Kommunikationsde-

The owner of the land where the airport is located is the municipality of Stockholm. The municipality has been letting it to the airport since 1936. The only political party in the municipality that during this period was consistently in favor of extending the contract beyond 2011 was the conservative party. In the 2006 election in Stockholm municipality the conservative party increased their number of seats by more than 40 percent, from 27 to 41 seats out of 101. This was the best result ever for the conservative party. The election outcome boosted the bargaining power of the conservative party in the negotiations with smaller parties in the center-right coalition. Rapidly, and behind closed doors, a new contract was negotiated. The new contract meant that operations would be extended to 2038. The decision on the operating contract was disclosed at a press conference in late September 2007. All the opposition parties issued minority reports before the council meeting where the new contract was debated, calling the process a coup. The news about the new contract was widely reported in local media.

4.2.2 Bromma Airport as a Negative Externality to Residents

The reason for the political controversy surrounding Bromma Airport is its geographic location. The airport is surrounded by residential housing and it generates a substantial negative externality on its surroundings - not the least in terms of noise. Figure 4.1 displays the noise contour from Bromma Airport as a dark red ellipse. The area inside the noise contour is subject to noise levels of 70 decibels

partementet, 1996). In 2000 Swedish Civil Aviation Administration presented a report on how they would phase out and eventually close Bromma airport by 2011 (Luftfartsverket, 2000).

(henceforth referred to as Lmax).[10] The contour is regarded as the best measure for defining the geographical area that is exposed to hazardous noise, as confirmed in a case in the Land and Environment Court (Miljööverdomstolen, 2010).[11] The contour is based on selected locations for noise measurement. The measurement data serves as input for interpolation and simulation in forming the contour. The measurement error of the border is +/- 100 meters.

Though houses within the noise contour are entitled to sound insulation according to the court ruling, Bromma Airport has permission to service take-offs and landings that generate more than 70 decibels between 6 am in the morning and 10 pm in the evening. Each aircraft type that operates on Bromma Airport must be tested and the upper bound is 89 decibels.[12]

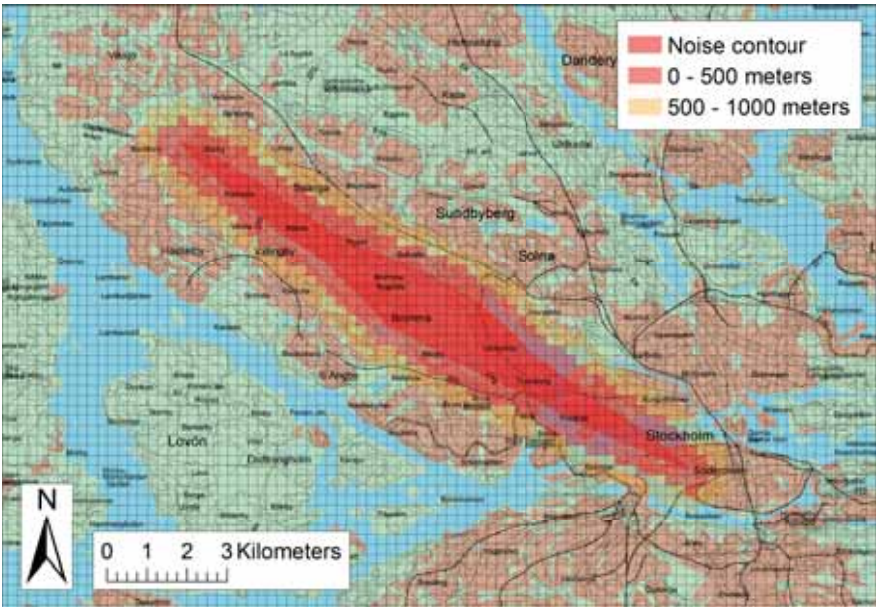
While noise is the obvious externality there are subtler negative externalities as well. Among those are transportation of fuel, the risk of accidents at take-off and landing, dumping of fuel etc. So, even though the established noise contour is the starting point for identifying geographical differences in negative externalities, we acknowledge that the magnitude of the negative externality is not perfectly measured and the noise contour of 70 decibels is only a proxy for exposure.

[10]The decibel scale is logarithmic. 60 decibels corresponds to a conversation in an office. 70 decibels are twice as loud. A vacuum cleaner is 70 decibels and makes it difficult to have a conversation. 90 decibels is four times as loud as 70 decibels and is the level of noise of a motorcycle at 8 meters distance. A Boeing 737 generates 97 decibels before landing at a distance of one nautical mile (1853 meters).

[11]The Land and Environment Court ruled that within this zone LFV, the regulatory body that oversees aviation in Sweden, must reimburse sound insulation.

[12]A noise level of 89 decibels is well above the threshold for when mandatory ear protection must be used in Swedish work places. In 2015 it was debated whether the noise calculations in the certificates had been manipulated and whether, in fact, some aircrafts generate more than 89 decibels. Thus, during day time the level of noise that some residents are exposed to may exceed 89 decibels.

Figure 4.1: Noise propagation around Bromma Airport



Note: The map shows the noise propagation around Bromma Airport along its runway which stretches from North-West to South-East. The dark red area is referred to as the noise contour. Within the noise contour the Land and Environment Court estimated that the noise level exceeded 70 decibels. The treatment region in our baseline specification is extended to also include the area which is less than 1,000 meters away from the noise contour. We call the treatment region the noise area. We locate house transactions and households on the grid. Each square is 250×250 meters. Source: Miljödomstolen (2006) and own analysis.

4.2.3 Definition of Treatment and Control Group

Noise transcends continuously and does not stop at the boundary of the contour. We define our baseline treatment group as the area that is exposed to 60 decibels or more. If the noise is generated from a

point source, it decreases by 6 decibels for every doubling of distance from the source. Let L_1 denote the decibel level at a distance of r_1 meters from the source. The noise level at a distance of r_2 meters is then given by L_2:

$$L_2 = L_1 + 10 \times \log^{10}\left(\frac{r_1^2}{r_2^2}\right) \tag{4.1}$$

This equation implies that if the boundary of the noise contour is located on average 500 meters from the source (i.e., a distance of two grid squares on the map in Figure 1), we should include residents located 1,000 meters further away in the baseline treatment group. This extended area, marked in yellow in Figure 1, is henceforth referred to as the noise area. Provided that the noise contour is approximated with an error of +/- 100 meters and household are located with some error (within a grid square) it is arguably a reasonable baseline definition.

The control group consists of homeowners in the municipality of Stockholm that are located further than 1000 meters away from the noise contour. In the following we will refer to real estate properties and households as being located either inside or outside the noise area. Later on we show that our estimates are robust to alternative definitions of the treatment and control group.

Based on the definition of treatment and control, we argue that the unexpected renewal of the airport contract in September 2007 was an exogenous shock to house prices for houses located inside the noise area. Thus, a standard difference-in-difference (DID) design can be used.[13] We return to details of the DID specification in Section 4.4.

[13]We are by no means the first to use the fact that a negative externality capitalizes into real estate prices. Chay and Greenstone (2005) and Currie et al. (2015) use it to measure the cost of nearby air pollution and toxic plants. See discussion in Section 4.5.1.

4.3 Data

We use two data sets in the analysis. First, we establish a house price effect of the contract renewal in a data set on transactions of property. Second, we analyze consumption responses in a household data set.

4.3.1 Data on Real Estate Transactions

Our transactions data is from the Swedish Land Surveying Agency. The Agency is responsible for registering title deeds, site-leasehold rights, and mortgage deeds. The data set covers all real estate transactions in the municipality of Stockholm over nine years; January 2004 to December 2012.

The data set includes the transaction price and the date of the transaction for each piece of real estate (20,588 transactions in total). We drop five transactions with price of zero and ten transactions with a price in excess of SEK 40 million which we believe are miscoded (there were no such expensive houses sold in Stockholm municipality during this time period). Further, we exclude pieces of real estate that we cannot be sure are intended to serve as a single family' primary residence.[14] This leaves 19,777 transactions. The data set also included the geographic location of the property in the form of GIS coordinates, the tax value, the area of the plot of land, and information about the houses, such as living area, supplementary area, age, if the land is owned or if it is long-leased and an index of the attractiveness of the location and the house standard. This index is used by the Swedish tax agency to assess the value of the property.

Table 4.1 reports summary statistics for these transactions. Panel A reports means and standard deviations of all transacted properties

[14]Pieces of real estate with a non-negative residential tax value

over the nine years. The average value per transaction is SEK 3.3 million and involves a single-family house with a living area of 117 square meters along with a total land area of 539 square meters. Panel B restricts the sample to transaction that took place up until 2008Q3 (i.e., up until the bankruptcy of Lehman Brothers). Since our aim is to isolate the effect of the contract's renewal in September 2007 and the financial crises had significant effects on the housing market in Stockholm, we will focus on this time period in much of our analysis. Panel C and D compare transactions inside and outside the noise area (within 1,000 meters away from the noise contour compared to further away). Real estate properties are essentially balanced before the renewal with two exceptions. Inside the noise zone, the buildings are older (52 years instead of 78 years) and the share of properties with site-leasehold rights (rather than outright ownership of the land through title deeds) is lower (12.5% versus 18.8%).[15]

4.3.2 Data on Households and Car Purchases

The household data set, provided by Statistics Sweden, covers all households whose residential address is located in the municipality of Stockholm. Within the municipality of Stockholm there are about 410,000 dwellings. Approximately 90% of these are apartments and 10% are single-family residential houses. We focus on households that own single-family houses since it is difficult to assign market values to apartments based on Swedish registries. Our sample consists of 41,147 households (corresponding to 112,581 individuals). Statistics

[15]Approximately five percent of all transactions involve real estate properties with a very low price. These are, most probably parceled plots of land that may not have an actual house on them. We choose to include them in our analysis of house prices since the effect on their market value is part of the households' wealth shock. Nevertheless, we henceforth refer to house price effects of the renewal despite a small share of properties that potentially only consist of land.

Table 4.1: Real estate transactions in the municipality of Stockholm (2004–2012)

	Price (kSEK)	Tax value (kSEK)	Age (Years)	Standard (Index)	Plot area (m^2)	Living area (m^2)	Non-living area (m^2)	Site-leasehold right (Y/N)
				A. Full sample (2004Q1-2012Q4)				
Mean	3351	1751	77.16	28.37	539.5	116.7	47.67	0.155
SD	2244	822.9	209.2	4.320	367.2	38.12	59.74	0.362
N	19,777	19,777	19,666	19,666	19,777	19,666	19,666	19,777
				B. Before financial crises (2004Q1-2008Q3)				
Mean	2947	1762	71.35	28.61	540.6	117.2	47.81	0.174
SD	1955	857.6	189.3	4.323	349.4	38.52	35.93	0.379
N	11,321	11,321	11,308	11,308	11,321	11,321	11,308	11,308
				C. Inside noise area and before renewal of contract (2004Q1-2007Q3)				
Mean	2807	1729	51.47	28.97	563.5	120.1	48.17	0.125
SD	1517	725.0	61.00	4.386	329.2	36.49	40.12	0.331
N	2330	2330	2329	2329	2330	2329	2329	2330
				D. Outside noise area and before renewal of contract (2004Q1-2007Q3)				
Mean	2672	1766	78.17	28.32	533.1	115.6	48.19	0.188
SD	1893	837.2	217.0	4.173	363.1	39.29	34.40	0.390
N	6926	6926	6926	6926	6926	6926	6926	6926

Note: All values in 1000s of Swedish kronor (SEK). The exchange rate is about 8 SEK/USD.

Sweden uses GIS coordinates to locate each household's residential address on the 250 times 250 meter grid illustrated in Figure 1. As in for instance Calvet et al. (2007) and Vestman (2017) we have detailed balance sheet information for each household up until 2007, that is before the renewal of Bromma Airport's contract.

For each household the market value of residential real estate is reported. However, in principle the property may be located somewhere else, e.g. if it is a summer house that for tax purposes is a property intended to serve as a permanent home. There are cases where the location of the residential address on the grid indicates that it is unlikely that the household owns the property in which it lives. We choose to exclude households which Statistics Sweden locates outside of Stockholm municipality or in the city center where there is no single-family residential housing. We also choose to exclude households where the primary property is untaxed because such properties are unlikely to be residential. We further exclude households that own several properties since we cannot infer where they are located on the grid. Lastly we restrict attention to households owning at least 50% of a property. After implementation of these restrictions it is reasonable to assume that households housing wealth refers to property which is located in the municipality of Stockholm and hence on the grid. We arrive at a sample of about 35,386 households owning a total of 32 037 houses.[16] Since the municipality of Stockholm municipality had a total of 41,000 single-family houses in 2006, we cover 78% of them.

Table 4.2 reports information about the entire population of households in our sample at the end of 2006 (column 1 and 2). It also reports the information divided into those outside the noise area (column 3 and 4) and those inside (column 5 and 6). Households inside the noise area have on average 2.88 members. Households outside

[16]A detailed description of sample restrictions can be found in Appendix A.

have marginally fewer (2.8). The average age of the oldest household member is 53 inside the noise area and 54 outside. In terms of wealth the two groups are very similar. The average financial wealth is SEK 959,000 compared to SEK 979,000. By Swedish standards these are substantial amounts of financial wealth. It indicates the sample of owners of single-family houses is a positively selected group. Based on Statistics Sweden's appraisal model for single-family houses both groups' average housing wealth is just over SEK 4.7 million.[17] We define the loan-to-value ratio as total debt divided by housing wealth. The average combined loan-to-value ratio (CLTV) is somewhat lower inside the noise area (0.37) than outside (0.42). These CLTVs are quite moderate and consistent with households being middle-aged and fairly wealthy. Housing wealth to net worth is also very similar for households inside and outside (1.50 versus 1.43). The high standard deviations indicate the non-linear nature of the measure. Gross labor income and disposable income (gross labor income plus transfers minus taxes) are also very similar and higher than the average Swedish households'.

In addition to demographic and balance sheet information we observe car ownership. We know the model of the car, the year it was produced, how many previous owners it has had, and the mileage. For each year and each car we see the last three transactions, the date it was bought, who sold it (given that they are in our sample) and who bought it. The data set includes the price if the car is new but not if it is used. Table 4.3 reports transaction statistics. In 2006, one in twenty households inside the noise area purchased a new car compared to one in twenty-five household outside the noise area.

[17]For each municipality, Statistics Sweden uses the ratio between observed transaction values and the tax value (which is based on the index for quality – see Table 1) to appraise the stock of single-family houses. Prior to the extension of the operating contract (2007Q3) this is in all likelihood an unbiased and quite accurate estimate of the market value.

Table 4.2: Households in single-family houses in Stockholm (2006)

	Full sample		Outside noise area		Inside noise area	
	Mean	SD	Mean	SD	Mean	SD
Distance from noise contour (meters)	3099	2436	4055	2088	314.8	338.9
Age	53.95	14.46	54.19	14.51	53.25	14.28
Household size	2.817	1.338	2.795	1.333	2.880	1.354
Financial wealth	974.5	2928	978.7	3042	959.2	2528
Housing wealth	3791	10884	3758	6171	3756	6846
Total wealth	4765	11671	4736	7263	4716	7615
Total debt	1256	1848	1253	1988	1264	1363
CLTV	0.407	2.752	0.421	3.168	0.365	0.567
Housing / Net worth	1.448	22.1	1.429	24.26	1.498	13.98
Labor income	570.7	580.3	567.4	606.6	579.6	494.1
Capital income	52.54	864.5	52.57	718.0	51.01	1185
N	35,388		26,410		8,978	

Note: All values in 1000s of Swedish kronor (SEK). The combined loan-to-value ratio (CLTV) and the housing-net worth ratio include all kinds of debt, including students loans and consumer credits. The exchange rate is about 8 SEK/USD.

Households inside are also somewhat more likely to purchase a used car. These statistics line up reasonably well with aggregate statistics for Sweden where national average for new (used) cars was about 6% (23%). That the likelihood of purchasing a car is somewhat lower in our sample is not surprising as households who live in a large city are less prone to own a car. Since we do not observe the price for used cars we will exploit characteristics of the used cars instead. Those characteristics are milage, age and whether the car belongs to a premium brand. We define three brands as premium; Audi, BMW, and Mercedes. In our sample, 2.8% of households buy a used car from one of these brands.

4.3.3 Measurement Issues

Overall, Swedish registry-based data is considered to be of very high quality. For our purpose there are two measurement issues that need to be addressed. The first issue is the date of the location. We observe where households live at the end of 2006 and use this location to form the treatment (noise) and control group. The announcement of the renewal of the contract, however, occurs at the end of the third quarter in 2007. In general there are not many movers in our sample (7% of our sample moved during 2007) so this classification error should not be large. As a robustness test we have excluded households that move during 2007 and the results remain the same.[18]

The second issue is that we observe ownership of the cars at the end of each year. At that time we see the three last transactions during the year for each car and the date of those transactions. This means that for cars that have changed ownership more than three times in a year we will not observe every purchaser. We will also not observe a household's purchase if the car is scrapped in the same year.

[18]There results are left untabulated but are available on request

Table 4.3: Car purchases per household: summary statistics

	Full sample		Control, before renewal		Treatment before renewal	
	Mean	SD	Mean	SD	Mean	SD
Car purchases	0.218	0.658	0.213	0.627	0.225	0.183
			A. New cars			
Yearly purchases	0.042	0.212	0.042	0.213	0.042	0.209
Price	247.8	131.0	249.8	138.6	242.9	101.4
			B. Used cars			
Yearly purchases	0.176	0.619	0.171	0.584	0.183	0.607
Milage (km)	12677	18156	13015	19118	13203	19825
Age (years)	7.026	5.484	7.015	5.464	7.059	5.538
Premium brand (0/1)	0.0278	0.200	0.0279	0.203	0.0275	0.176

Note: All values in 1000s of Swedish kronor (SEK). The exchange rate is about 8 SEK/USD. For 2005 and 2006 the price of new cars is based upon hand collection since Statistics Sweden only have price data for 2007 and later. The coverage for 2005 and 2006 is 2,398 out of 3,019 transactions (79,4%). Premium brands refer to Audi, BMW, and Mercedes.

Such events are not common and in any case we have no reason to suspect that they should correlate with the extension of the airport contract.

4.4 Empirical Strategy

4.4.1 Car Consumption and Housing Wealth

The population function of interest, if we were to relate changes in housing wealth to car consumption, could be postulated as:

$$\text{Car}_{it} = \alpha + \beta \log(\text{House price}_{it}) + \varepsilon_{it} \tag{4.2}$$

where Cars_{it} represents some measure of car purchases of household i in time period t. and $\log(\text{House price}_{it})$ is the house value of household i in time period t. The error term ε_{it} represents all other factors related to the outcome. Our baseline specification focuses on house price changes rather than changes to housing net worth (Mian et al., 2013). See Kaplan et al. (2016) for a discussion of alternative covariates and specifications.

4.4.2 First stage, Reduced Form, and IV-estimator

We define the treatment period to start on 1^{st} of October 2007, also denoted as 2007Q4. The variable NoiseArea_i defines the treatment group in the sense that the variable takes on a value of one if the household is located within 1,000 meters from the noise contour and zero otherwise. The variable Post_t is zero up until 2007Q4 and one thereafter. A standard difference-in-difference first stage equation reads:

$$\log(\text{House price}_{it}) = \gamma_0 + \theta\text{Noise area}_i + \delta(\text{Noise area}_i \times \text{Post}_t) + \gamma_t + \varepsilon_{it}^h \tag{4.3}$$

where γ_t indicates time effects and ε_{it}^h is an error term. Thus, θ measures the average (percentage) difference in house prices between the houses inside and outside the noise area. δ is the coefficient of interest and measures in the percentages change in house prices due to the extension of the contract. Throughout, standard errors are based on clustering of error terms at the level of the 250*250 meter grid.

The reduced form is identical to equation 4.3 with respect to the right-hand side but the outcome is our measure of car consumption as described in equation 4.2:

$$\text{Cars}_{it} = \phi_0 + \rho\text{Noise area}_i + \phi(\text{Noise area}_i \times \text{Post}_t) + \phi_t + \varepsilon_{it}^l \tag{4.4}$$

where ϕ_t indicates time effects, ε_{it}^l is an error term and ϕ is the parameter of interest. It measures the percentage change in purchase prices of new cars due to the renewal of the airport contract.

For δ and ϕ to have a causal interpretation the identifying assumption is the assumption of parallel trends prior to treatment between the treatment and control groups. Since equations (4.3) and (4.4) have causal interpretation only if that assumption holds, we will present dynamic estimates that allow for statistical hypothesis testing of the assumption. There should be no significant effects before the actual declaration of the extension in 2007Q3 (sometimes this is referred to as placebo tests). An immediate reaction after 2007Q3 adds further credibility to a casual interpretation.

4.4.3 IV Estimates of the Housing Wealth Effect

In order to obtain the instrumental variables effect we can weigh the reduced form with the first stage (i.e., $\hat{\beta}^{\text{IV}} = \hat{\phi}/\hat{\delta}$). This is a valid local average treatment effect, provided that the exclusion restriction holds. This second-stage IV estimates have the same interpretation as in our discussion of equation (4.2).

In practice, the effect of treatment on house prices is only observed in the data set of house transactions and not in the household data set. Therefore, we estimate the housing wealth effect using a two-sample instrumental variables approach analogous to Angrist and Krueger (1992). Inoue and Solon (2010) show that the inference problem is solved using a generated regressor correction as proposed by Murphy and Topel (2002). We follow the implementation of Fredriksson and Öckert (2013).[19]

4.4.4 Exclusion Restriction and Robustness

Our IV-estimator measures the causal effect of a change in housing wealth on car consumption if the exclusion restriction holds. Conversely, it does not measure the casual effect if the instrument (i.e., $\text{NoiseArea}_i \times \text{Post}_t$) has an independent effect on car consumption aside from the effect through housing wealth. In our baseline model our instrument is $\text{NoiseArea}_i \times \text{Post}_t$, which takes on the value one after 2007Q3 for the noise area and zero otherwise. Thus, a violation of the exclusion restriction amounts to no less than finding another event (e.g., another policy change) at the same time that affects car consumption in- and outside the noise area differentially. Clearly, building new houses on the land of the airport could create new public amenities. However, there are substantial existing pub-

[19]We thank Peter Fredriksson and Björn Öckert for sharing their Stata code with us.

lic amenities in the areas surrounding the airport and building new houses on the premises of the airport would take at least ten years and before that more public amenities (e.g., buses, schools and so on) would not be set in.

Further, we know that noise decreases with distance so we expect the treatment effect to monotonically decrease with the distance from the border of the noise contour. This would be compelling evidence of a causal effect as it is even more difficult to think of another event that would not occur in late 2007 and be monotonically related to the distance to the noise contour. We extend our regression model in two different ways to investigate this. First, we split our treatment group into three groups. The first group is defined as within the noise contour, the second is located up to 500 meters away from the border of the noise contour and the third between 500 and 1000 meters away from the border of the noise contour. We refer to this analysis as a dose response analysis. We would expect to find the smallest effect for the groups furthest away and the largest for the group within the noise contour. In addition, we consider a continuous measure of distance, denoted by $\log^{10}(\text{distance}_i^2)$ consistent with the propagation of noise given by equation (4.1).

We consider three additional robustness tests. Firstly, a standard robustness test to investigate compositional bias by adding controls for house characteristics. We consider the characteristics reported in Table 4.1. Secondly, we consider different coarser clustering schemes. In order to confirm that our way of clustering errors at the level of the 250 times 250 meter grid is innocuous we construct three coarser grids of 500 times 500, 1,000 times 1,000, and 2,000 times 2,000 meters, respectively. Statistical significance in key estimates are not altered. Finally, difference-in-difference estimates are known to be sensitive to functional form. In our baseline specification we use the natural logarithm of prices. However, we also consider a linear specification.

In general we obtain very similar results.

Since our model is robust to the above extensions and shows clear signs of monotonically decreasing dose response, we argue that our results are likely to be causal and claim that it is extremely hard to understand what other type of other channel could violate the exclusion restriction.

In a difference-in-difference approach one might also worry about general equilibrium effects. In our case these effects are likely to be small since 90% of households in Stockholm municipality reside in apartments as opposed to single-family houses. Moreover, only 17% of households residing in single-family houses in Stockholm live inside the noise area. Further, the self-contained houses in Stockholm municipality only constitute about 16 % of the total number of self-contained houses in the larger Stockholm area.

4.5 Results

This section reports our results. Firstly, we establish a robust effect on house prices in the housing transaction data set. Secondly, we validate that this effect is also present in the household data set. Thirdly, we present results on car purchases in the form of reduced form estimates as well as IV estimates. Finally, we then consider the role of initial leverage.

4.5.1 Effect on House Prices (First Stage)

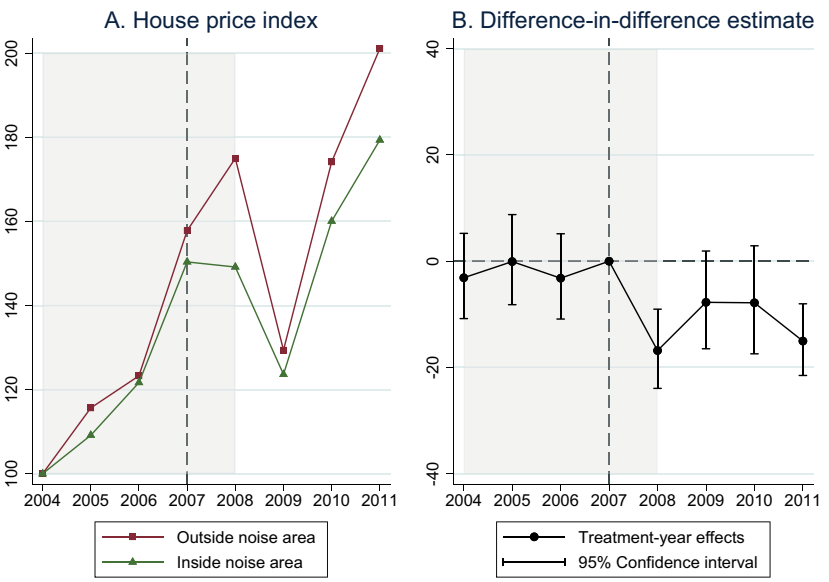
To validate our DID specification, we test for parallel trends in outcomes prior to 2007Q3. We augment equations (4.3) and (4.4) with yearly time dummy variables. We define a set of dummy variables based on time relative to treatment. Since the renewal of the operating contract was disclosed in late September 2007 we define every

12-month period as running from October 1 of year $t-1$ to September 30 of year t. We interact this set of time dummy variables with NoiseArea_i and omit the dummy variable for 2006Q4–2007Q3 so that it serves as the reference level for prices outside noise area.[20] That is, for every 12-month period, we estimate a treatment effect relative to the 12 months just before contract renewal. This is the standard test of parallel trends (Angrist and Pischke, 2009).

Moreover, this specification enables illustration of the dynamics after 2007Q3 (i.e., post contract renewal). We start by evaluating house prices (i.e., the first stage) in Figure 4.2. Panel A displays house price fluctuations inside and outside the noise area. The price series are indexed to 100 in 2004. The series follow the same trend from 2004Q1 to 2007Q3. Subsequently, there is a gap between them. Panel B displays the corresponding DID estimates on log house prices. Common pre-trends up to 2007Q3 cannot be rejected. However, the immediate effect in 2007Q4-2008Q3 is -16.7 log points. Differences remain throughout up until 2011. However, since our aim is to isolate the effect of the contract's renewal in September 2007 we focus on the time period up until the major onset of the financial crisis in our main specifications, that is from 2004Q1 to 2008Q3. In Figure 4.2 and in all subsequent figures this area is shaded in grey. The focus on a short post-treatment time window should come at little cost from a causal point of view since a DID strategy is more credible if effects appear quickly. Appendix B discusses several noteworthy aspects of the housing market, such as how the house price index construction method underlying Panel A compares to official providers' index methods, an investigation (and rejection) of compositional effects of transacted properties, and how the culmination of the financial crisis affected the housing market.

[20] In our graphical illustrations, 2006Q4-2007Q3 is referred to as 2007 and so forth unless explicitly stated otherwise.

Figure 4.2: Effect on house prices



Note: Panel A shows house price indices outside and inside the noise area. Panel B shows the corresponding difference-in-difference estimates in log points. The timing on the horizontal axis is shifted one quarter. That is, 2007 refers to 2006Q4–2007Q3 and so forth. The shaded grey area indices the main sample period (2004Q1-2008Q3).

Table 4.4 reports estimates of the first stage. Columns (1) to (5) limit the sample to the pre-financial crisis period. Column (1) considers an unrestricted sample with no control variables and the estimated effect is −26% inside the noise area compared to outside. Column (2) restricts the sample to transactions of real estate for which the tax value is non-missing (which we interpret as excluding properties that are not intended to serve as single family housing). The estimate is reduced somewhat to around −21%. Column (3) considers a specification with basic control variables on property characteristics. We consider this to be our preferred specification since it controls for characteristics that serve as input to official index producers' methods. In this specification the magnitude of the effect decreases to −16.7% but remains highly statistically significant. Column (4) adds additional control variables, such as the age of the house, plot area, and the change in property tax due to the new property tax code coming into effect in 2008. These additional control variables affect the point estimate by only 2 percentage points. The fifth column considers a shift of the dependent variable from the log of the transaction price to the transaction price. The house price effect amounts to −419 kSEK, or 15% of the pre-treatment mean inside the noise area. Finally, column (6) considers the full sample period (i.e., 2004-2012). The long-term effect is muted because of the turbulence in the housing market but remains large and statistically significant. We conclude that our instrument is strong and that the effect on house prices inside the noise are is robust across specifications and sample periods.

Dose Response in House Prices

We provide further credibility to the casual mechanism behind the decrease in house prices close to Bromma Airport by considering dose responses. That is, we model the effect, or intensity of treatment, as

Table 4.4: Effect on house prices

	(1)	(2)	(3)	(4)	(5)	(6)
Noise area$_i$ × Post$_t$	-0.261***	-0.212***	-0.167***	-0.147***	-419.2***	-0.125***
	(0.051)	(0.039)	(0.038)	(0.032)	(120.0)	(0.026)
Noise area$_i$	0.190***	0.196***	0.081	0.087**	0.190***	0.114***
	(0.043)	(0.043)	(0.055)	(0.043)	(0.043)	(0.039)
N	11,696	11,321	11,321	11,308	11,321	19,777
R-squared	0.099	0.102	0.525	0.559	0.540	0.537
2004Q1-2008Q3	Yes	Yes	Yes	Yes	Yes	No
Restricted sample	No	Yes	Yes	Yes	Yes	Yes
Basic controls	No	No	Yes	Yes	Yes	Yes
Additional controls	No	No	No	Yes	No	No
Linear specification	No	No	No	No	Yes	No

Note: The restricted sample only includes transactions of real estate that have a non-missing tax value. The basic controls include the tax value, an indicator whether the property is a site-leasehold right, and an indicator for a low sales price (<250 kSEK). The set of additional control variables include age, standard, plot area, living area, and non-living area, and the change in amount of property tax to be paid due to the tax reform coming into effect in 2008. The specification in column (3) corresponds to Panel B of Figure 4.2. Clustering of standard errors at the level of the 250×250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

a function of the distance to the noise contour. In terms of empirical methods, this brings our analysis close to the literature that uses quasi-experimental methods to quantify the negative externalities of local changes to environmental conditions (e.g., Currie et al., 2015).[21]

Column (1) of Table 4.5 presents results from dividing the treatment effect into three separate effects based on the distance from the noise contour. The first covariate, $\text{Post}_t \times 0$, measures the effect of the renewal inside the noise contour in comparison to house prices further away than 1,000 meters from the border of the noise contour. Notably, the effect is of a large magnitude; -22%. The second covariate, $(0 \quad 500] \times \text{Post}_{it}$, captures the effect on prices just outside the noise contour but not more than 500 meters away. The effect in this geographical region is smaller (-16%). The effect is further muted another 500 meters away from the noise contour (-11%). The monotone relationship (i.e., a dose response) is compelling evidence in favor of a casual interpretation of a price effect from the contract's renewal. Going forward, we refer to this measure of treatment as "segmented distance". Column (2) considers a continuous measure of treatment rather than a discrete one. Assuming that equation (4.1) is a good proxy for the net effect of all kinds of externalities (i.e., not only noise), the effect on house prices should be proportional to the logarithm of distance (or distance squared). We interact this intensity of treatment with the post contract renewal indicator. The point estimate (0.120) is statistically significant. Going forward, we refer to this measure as "log distance". The remaining five columns, (3) to (7), rely on specifications similar to the baseline but with different definitions of the noice zone. The baseline model defines the noise area as the area which is up to 1,000 meters away from the noise con-

[21] The specifications of this section rely on more continuous measures of distance. Panel A of Figure 4.12 in the Appendix E confirms that transacted real estate properties are evenly distributed up to 10 kilometers away from the noise contour.

Table 4.5: Effect on house prices: dose response

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Post}_t \times 0$	-0.229***						
	(0.036)						
$\text{Post}_t \times (0 \quad 500]$	-0.162***						
	(0.077)						
$\text{Post}_t \times (500 \quad 1000]$	-0.111***						
	(0.046)						
$\text{Post}_t \times \log^{10}(\text{dist}^2)$		0.120***					
		(0.024)					
$\text{Post}_t \times 0$			-0.204***				
			(0.036)				
$\text{Post}_t \times [0 \quad 200]$				-0.226***			
				(0.050)			
$\text{Post}_t \times [0 \quad 400]$					-0.212***		
					(0.044)		
$\text{Post}_t \times [0 \quad 600]$						-0.173***	
						(0.042)	
$\text{Post}_t \times [0 \quad 800]$							-0.164***
							(0.039)
N	11,321	11,321	11,321	11,321	11,321	11,321	11,321
R-squared	0.526	0.53	0.524	0.559	0.525	0.524	0.525
2004Q1-2008Q3	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Restricted sample	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Basic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports dose responses as Post_t is interacted with variables that indicate the distance to the noise contour. A distance of 0 indicates inside the the noise contour. The noise area includes the noise contour plus up to 1,000 meters. The restricted sample only includes transactions of real estate that have a non-missing tax value. The basic controls include the tax value, an indicator whether the property is a site-leasehold right, and an indicator for a low sales price (<250 kSEK). Clustering of standard errors at the level of the 250 × 250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

tour. In Column (3), we reduce this to 800 meters away, In Column (4), 600 meters away and so on to Column (7), where we define the noise area to be the noice contour. In all of the specifications the effect is large and statistically significant, illustrating that the analysis is not sensitive to a particular definition of the treatment and control groups. Further, the effect decreases monotonically as the area widens, consistent with notion of a dose response. Taken together, Table 4.5 is evidence of a distinct effect on house prices, regardless of exactly how the treatment group is defined.

Table 4.10 in Appendix C illustrates that the estimates are robust to using alternative sub-sets of the control group, providing additional evidence for tight identification.[22]

4.5.2 A Cross-Walk Between the Real Estate Transaction Data Set and the Household Data Set

We now turn to consumption responses in the household data set. As a prologue to this analysis we establish that the effect on house prices of the contract renewal in September 2007 is present in the household data set too. As previously explained, the fundamental obstacle is that Statistics Sweden's model for appraising single-family houses in the household data set is too coarse, in the geographical sense, to be able pick up the treatment effect of the renewal of the airport contract.[23] That said, the house price effect ought to be present as a relative loss in realized capital gains from 2007Q4 and onwards for house sellers. In a DID regression with years 2004-2006 and 2008

[22]In untabulated results we have confirmed that changing the level of clustering to 500×500, 1,000×1,000 or 2,000×2,000 standard errors does not affect the level of statistical significance in any substantial way.

[23]To be precise, Statistics Sweden employs the appraisal method municipality by municipality.

we find a highly statistically significant decrease in realized capital gains inside the noise area, amounting to SEK 28,900 per household and year.[24] This is the intent-to-treat (ITT) effect measured over all households inside the noise area, of which most do not sell. Our transaction data set suggests that there are about 2,400 transactions every year among the stock of 40,000 single-family houses in the Stockholm municipality which implies a rate of sales of 6%. After scaling we arrive at approximate estimate of a loss of 482 kSEK. This is very close to the estimates of Table 4.4. Figure 4.3 reports the same finding graphically. We conclude that the estimated ITT effect in realized capital gains in the household data set confirms our findings on house price effects in the housing transaction data set.

In contrast to the effect on capital income, we find no effect on other kinds of income such as labor income, total income from employment, labor income plus unemployment benefits, or disposable income, (see Appendix E (Figures 4.13 and 4.14)). This corroborates our view that the renewal of Bromma Airport's contract had no major economic consequences or that that the local house price shock in itself affected general economic conditions.[25]

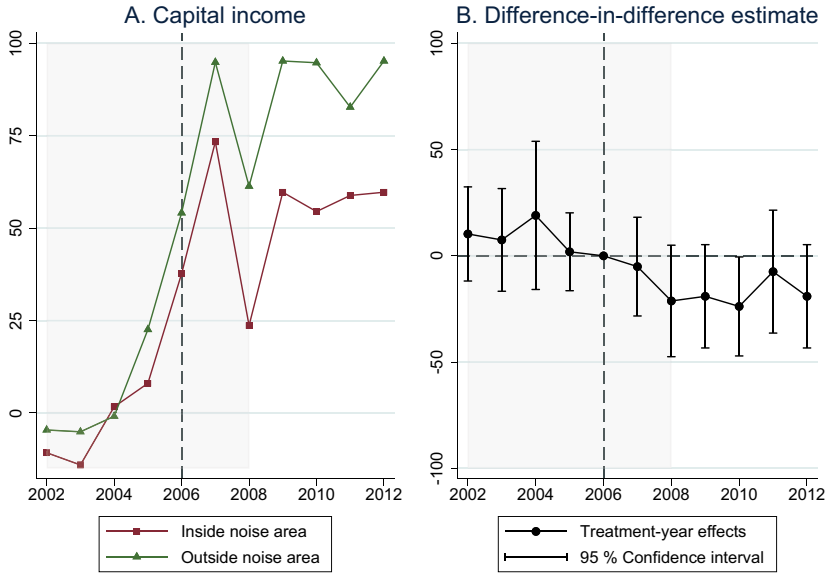
4.5.3 Effect on Car Purchases (Reduced Form)

We briefly consider reduced form estimates on car consumption, as specified by equation (4.4), before moving to the IV estimates.

[24]We exclude 2007 from the regression as is partly treated and we only have data on capital income on a yearly basis.

[25]While we do not challenge the view that airports are important for economic activity it would be surprising if Bromma Airport would be of vital importance to Stockholm. It is a considerably smaller airport than Arlanda Airport which is located only 30 kilometers away. Furthermore, according to wikipedia Bromma Airport itself only have 110 full-time employees, implying that it is not a major employer in Stockholm (https://sv.wikipedia.org/wiki/Stockholm-Bromma_flygplats).

Figure 4.3: Effect on capital income



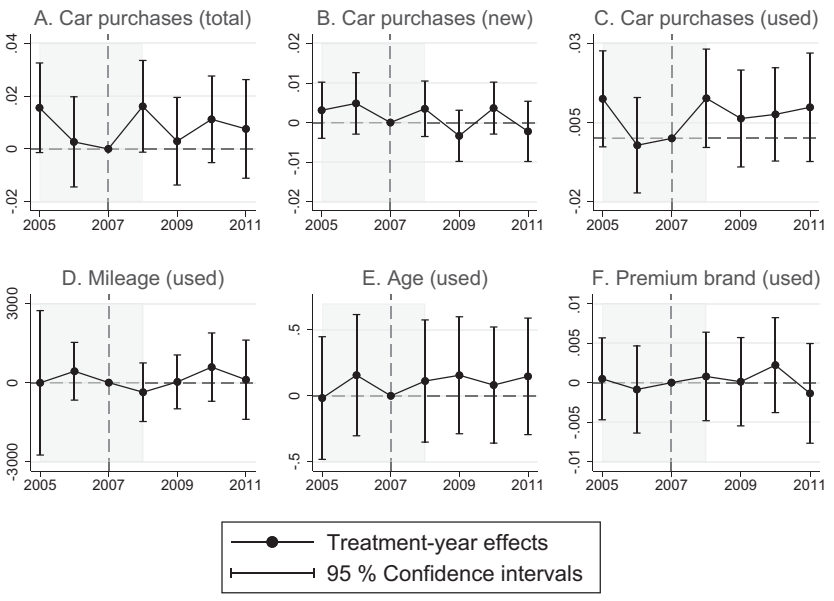
Note: The figure shows the capital income for households inside and outside of the noise area (Panel A) and the corresponding DID estimates (Panel B). The timing on the horizontal axis is calendar year. There are no control variables in the underlying regression.

Figure 4.4 summarizes our findings on extensive margin effects (i.e., the likelihood that a car is purchased) and substitution effects between different kinds of used cars. We find no evidence of housing wealth effects on these margins. There is no effect on the likelihood of buying any kind of car (Panel A), a new car (Panel B), or a used car (Panel C). If anything, point estimates are in the positive range. We can therefore rule out any negative effect on new cars in 2007Q4-2008Q3. Furthermore, we find no evidence that buyers of used cars inside the noise area substitute to cheaper cars because there is no effect on mileage (Panel D), age (Panel E), or on the likelihood of buying a used car with a premium brand (Panel F). Appendix D (Table 4.11) reports further details on the regressions corresponding to Figure 4.4.

As a consequence of the null effects on these margins we focus on analyzing the substitution effect among new cars for which we observe the price. Figure 4.5 reports prices of new cars. Again, the assumption of parallel trends prior to contract renewal holds up very well. There are no statistically significant effects before the renewal and the point estimates show no systematic pattern. After the renewal there is a distinct and statistically significant negative effect inside the noise area. Relative to the reference period, the average decrease in car prices inside the noise area is approximately -7%. After 2008Q3 the effect is somewhat muted by the financial crisis but the point estimate remains quite stable.

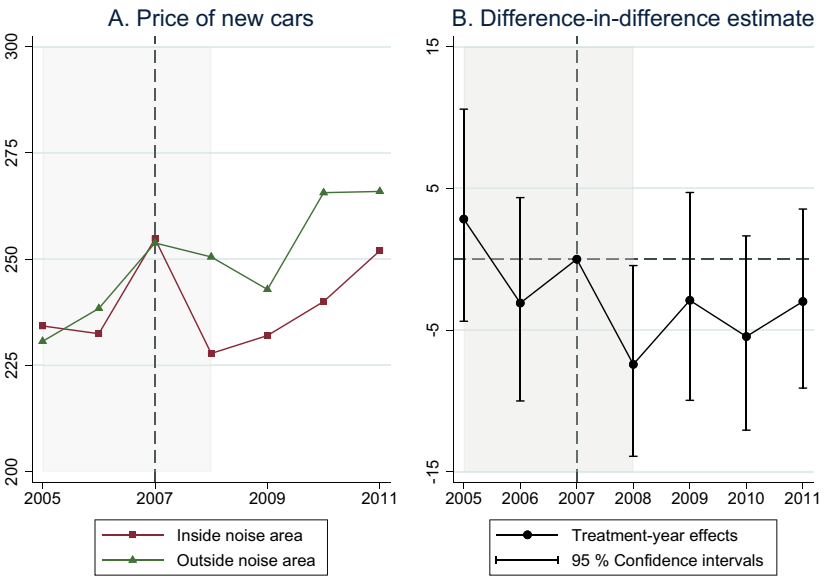
Figure 4.5 is based on a regression specification with no controls. Table 4.6 reports estimates of the substitution effect for new cars based on alternative specifications. Column (1) reports the equivalent of the figure but restricted to the primary time period of analysis. The effect on on prices of new cars is -7.2%. Column (2) corresponds to an addition of control variables that are likely to affect the preference for a new car. The control variables are; age of household

Figure 4.4: Car purchases: extensive margin and substitution effects



Note: The figure shows estimates of 12-months effects inside the noise area relative to outside the noise area. 2006Q4-2007Q3 is omitted to serve as reference year (the vertical dashed line). The timing on the horizontal axis is shifted one quarter. That is, 2007 refers to 2006Q4–2007Q3 and so forth. Premium brands refer to Audi, BMW, and Mercedes. The shaded grey area indices the main sample period (2004Q1-2008Q3).

Figure 4.5: Car purchases: price of new cars



Note: The figure shows estimates of 12-months effects inside the noise area relative to outside the noise area. 2006Q4-2007Q3 is omitted to serve as reference year (the vertical dashed line). The timing on the horizontal axis is shifted one quarter. That is, 2007 refers to 2006Q4–2007Q3 and so forth. The shaded grey area indices the main sample period (2004Q1-2008Q3).

head, household size, financial wealth, housing wealth, total debt, labor income, and capital income. At −6.0% the point estimate is essentially unaffected. Column (3) considers the price amount as a dependent variable rather than the log of the price. The estimate is −22 kSEK and is the equivalent of −9.1% given a mean price of 243 kSEK inside the noise area for the period 2005Q1-2007Q3. Finally, column (4) displays the estimated effect for a longer sample period, stretching up to 2011Q4. At −5.7% the effect is somewhat smaller but still statistically significant.

Table 4.6: Effects on the price of new cars (intensive margin)

	(1)	(2)	(3)	(4)
Noise area$_i$ × Post$_t$	-0.072**	-0.060**	-22**	-0.057**
	(0.029)	(0.029)	(8.7)	(0.025)
Noise area$_i$	-0.001	-0.001	-0.8	-0.017
	(0.017)	(0.029)	(5.4)	(0.013)
N	4,794	4,794	4,794	8,638
R-squared	0.010	0.077	0.076	0.015
Sample: 2005Q1-2008Q3	Yes	Yes	Yes	No
Controls	No	Yes	Yes	No
Linear specification	No	No	Yes	No

Note: The set of control variables includes age of household head, household size, financial wealth, housing wealth, total debt, labor income, and capital income. Clustering of standard errors at the level of the 250 × 250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We conducted additional analysis that favors the interpretation of a direct casual effect from house prices to responses in car purchases.

Analogous to the dose response analysis on house prices, we have investigated responses on new car purchases based on distance to the noise contour. The evidence is compelling in that the segmented distance displays a similar monotonically decreasing relationship with distance to the contour as the effects on house prices did and that the coefficient in the continuous specification is significant. Furthermore, statistically significant estimates do not rely on our baseline definition of the noise area. See Appendix D (Table 4.12).[26]

4.5.4 IV Estimates of the Housing Wealth Effect

We shift to reporting IV estimates of the elasticity between house prices and consumption as specified by equation (4.2).[27] We have established that there is no effect for the extensive margin and no substitution effect for used cars so focus is on the substitution effect for new cars.

Table 4.7 reports IV estimates for the subset of households that purchase a new car. Columns (1) and (2) report elasticities based on the baseline instrument.[28] Column (1) reports an elasticity of 0.505 which is statistically significant at the 1%-level. Adding control variables does not affect the estimate.

[26]We also ensured ourselves that the geographical distribution of car purchasers is similar to the distributions of real estate transactions and households – see Panels A to C of Figure 4.12.

[27]Our implementation of the IV estimator is as follows. We first estimate equation (4.3) in the data set on real estate transactions. We then use the estimates to predict the quasi-random variation in house prices in the household data set. Different functional forms of the predictive variables are considered. The main specification utiliizes NoiseArea_i and $\text{NoiseArea}_i \times \text{Post}_t$, but we also consider the specifications of columns (1) and (2) of Table 4.5. The predicted house prices are used as covariate in equation (4.2). The IV estimates may deviate somewhat from $\hat{\phi}/\hat{\delta}$ if the set of control variables in the two specifications differ.

[28]The baseline instrument is Noise area_i + Noise $\text{area}_i \times \text{Post}_t$. We correct standard errors for pre-estimation as previously described.

Table 4.7: Housing wealth elasticities

	(1)	(2)	(3)	(4)	(5)	(6)
log(house price)	0.505***	0.478***	0.421**	0.356**	0.327**	0.287**
	(0.192)	(0.211)	(0.189)	(0.182)	(0.143)	(0.139)
N	4,794	4,794	4,794	4,794	4,794	4,794
R-squared	0.011	0.078	0.011	0.078	0.010	0.077
Instrument	Noise area	Noise area	Segmented distance	Segmented distance	$\log^{10}$	$\log^{10}$
2005Q1-2008Q3	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes

Note: The table presents second-stage two-sample IV estimates using different functional forms of the instrument. Standard errors are corrected for first-stage estimation. The set of control variables include tax value, age, standard, plot area, living area, and non-living area and a control for the change in the property tax code in 2007 in the first stage and the age of household head, household size, financial wealth, housing wealth, total debt, labor income, and capital income in the reduced form. Clustering of standard errors at the level of the 250×250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We also consider using the instrument of column (1) in Table 4.5 which picks up finer variation in house prices (i.e., segmented distance). Column (3) reports the corresponding IV estimate of 0.421. It is somewhat smaller than the baseline version of the instrument. Adding control variables (column (4)) leads to some further reduction, however, the estimates remain highly statistically significant. Finally, we consider the continuous measure of the instrument, corresponding to column (2) of Table 4.5 (i.e., log distance). The point estimate is 0.327 and addition of control variables reduces it a bit further to 0.287. Nevertheless, it remains statistically significant. We conclude that the elasticity is in the range from from 0.3 to 0.5.

We convert the elasticities to a marginal propensity to consume out of housing wealth. The percent loss in housing wealth in the treatment group equals 16.7% which amounts to 469 kSEK ($0.167 \times 2,807$ kSEK). Based on an elasticity of 0.5, the reduction in spending amounts to 20 kSEK ($0.167 \times 0.50 \times 243$ kSEK). This implies that the MPC on new cars out of housing wealth is 0.043, or 4.3 cents per dollar. At face value, this implies that households which purchase a new car respond substantially. However, comparisons to estimates based on aggregate data require conversions as discussed below.[29]

4.5.5 Comparison to Previous Literature

To compare with estimates in the previous literature, different kinds of adjustments and considerations need to be made. Appendix F presents a summary table of estimates in the previous literature.

[29] At a conceptual level, one could study the difference between the purchase price of the household's new car and the sales price of the household's old car. For some theoretical considerations, this would be the more relevant decision margin. This is however not the typical framing in the empirical literature and from the perspective of aggregate consumption, such transactions of used cars net out. Nevertheless, we have not found any evidence of an effect on households' car sales, see Appendix D (Figure 4.11).

Comparison to Estimates Based on Aggregated Car Data

To make a comparison with other estimates of the MPC on cars based on regional data sets (e.g., aggregated to zip codes as in the case of the R.L. Polk data set for the USA), our estimate must be adjusted for the share of households that purchase a new car. We multiply our estimate with that share, which is 4.2% (see Table 4.3). Thus, the implied aggregate MPC based on our design is 0.0018, or 0.18 cents per dollar. Notice that this is a pure intensive margin response since we found no evidence of an extensive margin effect. Relative to Mian et al. (2013) our MPC is a order of magnitude smaller.[30] We discuss potential reasons for this below.

Comparison to Estimates Based on Total Household Expenditure

Most studies estimate elasticities between housing wealth and total household expenses. A comparison with these studies' estimates of MPC on total expenses require further assumptions on the allocation of (marginal) expenses between cars and other goods and services when exposed to a housing wealth shock. To the best of our knowledge, Mian et al. (2013) is the only study that both reports MPCs on cars and other goods. They find that the MPC on cars constitutes 42% of the MPC of total expenses per geographical unit. Dividing this share with our implied MPC on total expenses amounts to a mere 0.0044, or 0.44 cents per dollar. Alternatively, one can employ the

[30]They estimate the MPC to 0.023, or 2.3 cents per dollar in geographically aggregated data. The R.L. Polk data set reports the number of new automobile registrations. To arrive at an imputed value, the census expenditure weight is imposed. Consequently, they find a sizable effect driven by the extensive margin whereas we find a zero effect on that margin and many times smaller effect on the intensive margin. Another noteworthy difference is that we focus on homeowners exclusively whereas the aggregation to ZIP code level means that homeowners and renters are mixed in the analysis.

fact that 6.1% of Swedish household consumption is allocated to cars according to the national accounts.[31] If marginal adjustments are no different than average consumption allocations, the implied MPC on total expenses is 0.03, or 3 cents per dollar. It is well-known that the volatility of durable goods in general (and cars in particular) display a high time-series volatility, see Appendix E (Figure 4.15). Therefore, we view this as a definitive and unrealistic upper bound. Since the time series standard deviations of total consumption and car consumption differ by more than a factor of five, an MPC on total household consumption well below 0.006 is more realistic. Nevertheless, recent U.S. estimates are considerably larger. Appendix F summarizes the literature and finds estimates in the range from 0.01 to 0.06.

Differences to Previous Studies

One may ask what a reasonable theoretical housing wealth effect is. If the housing wealth shock is viewed as permanent and the mortgage and credit markets are well-functioning, if households are not borrowing constrained after the shock, and if the mortgage cost is unaffected, then the reduced wealth effect should be smoothed during the remainder of life. If households cut back equally on all goods and services, the total adjustment on cars should be 29.9 kSEK (6.1%×469 kSEK). We reported that 21% of households purchase a car every year which can be translated to once every fifth year on average. Remaining life is approximately 30 years in our sample, so households in our sample will buy six more cars. Assuming that the optimal consumption path is flat, each remaining car purchase should then be 4.8 kSEK less expensive. Assuming that new cars constitute 50% of the total expenditure on cars combined, and the fact that we only observe ef-

[31]Source: Statistics Sweden, Table NR0103B0.

fect on new cars, this would imply an MPC of 0.02, or 2 cents per dollar, among households that purchase a new car. This is half the size of our estimate. Nevertheless, it is within the 95%-confidence interval of our baseline estimate and closer than that to the estimates that are based on the alternative instrument specifications.[32]

This said, we turn to an important discussion of contextual differences. One obvious difference is that house prices typically fall as a consequence of changing macroeconomic conditions, such as lower income or more negative views of the future. Most U.S. studies therefore rely on geographical (topological) variation of Saiz (2010) to instrument for housing supply. In a recent leading study, Aladangady (2017) interacts this variation with the long-term interest rate to obtain relevant variation in the financing cost of housing. Nevertheless, it is possible that the exclusion restriction is invalid in that the economic activity in regions with different topological characteristics is affected differently to long-run interest changes in other ways than through impact on house prices. Davidoff (2013, 2016) has raised concerns about the validity of the exclusion restriction. While the response of households in such an economic environment may be what we ultimately need to consider in economic policy making and in preemptive macroprudential regulatory work, our estimate is cleaner in that there are no contemporaneous general equilibrium effects.[33] Absence of any general equilibrium effects should it make it easier to

[32]This back-of-the-envelope calculation of course disregards preference heterogeneity between purchasers of new and used cars, different MPCs across different goods, credit constraints, intergenerational transfers and that the optimal consumption profile may not be flat.

[33]In order to generate as large house price swings as in the Great Recession (i.e., comparable to the effect near the airport), Kaplan et al. (2017) consider time-varying beliefs about future housing demand jointly with changing credit conditions. Our source of house price variation have no economy-wide implications.

use our estimate in the calibration of a structural model.[34]

Another noteworthy difference is that many U.S. estimates include data from the Great Recession. Thus, one important difference may be that our estimate is a "good times" estimate. As is evident from Panel A of Figure 4.2, households close to the airport did not suffer a accual wealth loss. Their housing wealth did just not appreciate at the rate as households in the rest of Stockholm. If households have a concave utility function this matters. In essence, we estimate the MPC out of households located outside the noise area and they respond little in the face of a house price increase. In contrast, in a design where one group suffers from a house price decrease the MPC would be slightly greater. Nevertheless, estimates would not differ by an order of magnitude.[35]

4.5.6 Effects by Loan-to-Value Ratio

A central concern in the housing wealth literature is how to distinguish a pure housing wealth effect from the effect of a binding collateral constraint. DeFusco (2018) reports sizeable MPC estimates of relaxing the borrowing constraint, holding expected wealth constant. Also Cooper (2013) finds a strong collateral effect. Relatedly, Benmelech et al. (2017) find substantial effect of a credit crunch in the asset-backed commercial paper market on auto purchases during the financial crisis, attributing it to tighter credit supply.[36]

We investigate if there are heterogenous responses among households depending on pre-treatment balance sheet characteristics, in particular whether households with different (combined) loan-to-value

[34]See Berger and Vavra (2015) and Berger et al. (2017) for seminal contributions.

[35]See page 3420 of Aladangady (2017) for a discussion.

[36]Laerkholm Jensen and Johannesen (2017) find effects of tighter credit supply in Danish data.

ratios (CLTVs) respond differently. We divide our sample into subgroups based on their LTV at the end of 2006 (i.e., prior to treatment). We create three groups based on whether it is likely that a 17-percentage-point loss in housing wealth is likely to make a household borrowing constrained. The first group is one with low CLTVs which we define as being less than 0.6. The second group is the medium one with CLTVs between 0.60 and 0.80 and the third one consists of households with high CLTVs greater than 0.80. Households in the low group are unlikely to face binding constraints even in the presence of a 17-percentage point price shock whereas households in the high group are likely to face meagre possibilities to obtain additional mortgage loans using their house as collateral even before the house price shock.

The cut-off values are thus chosen so that it is likely that the collateral constraint becomes binding for the middle group. For households in this group a house price fall of 17% would put them close to or at their borrowing constraint. We would therefore expect to see changes in behavior in this group if housing wealth is monetized through credit to finance purchases of new cars.

Panel A of Table 4.8 reports the reduced form estimates. Columns (1) to (3) illustrate that the negative effect of the house price shock is entirely concentrated to the medium group. The magnitude of the estimate for that group is four times as large as the previously reported baseline estimates. Column (4) considers an interaction with a linear and squared term, the latter being highly statistically significant and suggesting a maximum response of -0.10 at a CLTV ratio of 0.5 but a zero effect at CLTVs close to 0 and 1. Turning to IV estimates in Panel B, the house price elasticity is estimated to 1.17 in the medium group and column (4) confirms a concave relationship with implied elasticity in the range from 0 to 0.50. We conclude that the initial CLTV is critical for households' response. Importantly, a

Table 4.8: Effects by loan-to-value ratio

	(1)	(2)	(3)	(4)
A. Reduced form estimates				
Noise area$_i$ × Post$_t$	-0.044	-0.234***	0.087	-0.005
	(0.033)	(0.088)	(0.103)	(0.057)
Noise area$_i$ × Post$_t$× CLTV				-0.383*
				(0.227)
Noise area$_i$ × Post$_t$× CLTV^2				0.377**
				(0.193)
B. IV estimates				
log(house price)	0.223	1.174**	-0.441	0.026
	(0.175)	(0.516)	(0.524)	(0.294)
log(house price) × CLTV				1.997*
				(1.153)
log(house price)× CLTV^2				-1.937**
				(0.964)
N	3,765	546	483	4,794
Restriction on the CLTV ratio	0.0-0.6	0.6-0.8	0.8>	No
2005Q1-2008Q3	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Note: The table presents second-stage two-sample IV estimates. Standard errors are corrected for first-stage estimation. The set of control variables include tax value, age, standard, plot area, living area, and non-living area and a control for the change in the property tax code in 2007 in the first stage and the age of household head, household size, financial wealth, housing wealth, total debt, labor income, and capital income in the reduced form. Clustering of standard errors at the level of the 250 × 250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

high CLTV may, in fact, mute the short-term response to a house price shock.

4.6 Concluding Remarks

Long before the financial crisis economists understood that the housing market is of importance for the evolution of the macro economy, for instance through the wealth effect it exposes homeowners to. However, because of weak identification, the credibility of available estimates of the elasticity between house prices and household consumption has been low.

We base our estimate on a quasi-experiment that occurred in Stockholm, Sweden, in September 2007. As a result of political bargaining behind closed doors it was decided that Bromma Airport, located in a residential area, would not end its operations in 2011. Rather, its operating contract with the municipality was renewed and operations were allowed to continue until 2036. Using a data set of all transactions of single-family houses, difference-in-difference estimates indicate that prices of houses located close to the airport's noise contour fell by approximately 17% soon after the renewal of the operating contract. The closer to the airport's noise contour the greater impact.

Since the shock to house prices was local, unanticipated, and uncorrelated with changing macroeconomic conditions it provides a unique opportunity to study the housing wealth effect. We estimate the elasticity between house prices and consumption, using a population data set of all homeowners that live in single-family houses in Stockholm and a two-sample instrumental variables approach. We find a sizeable MPC of 4.3 cents per dollar among households that purchase a new car. It is concentrated to car purchasers that have a combined LTV ratio between 0.6 and 0.8. For this group, the re-

sponse amounts to an MPC of more than 10 cents per dollar.

Based on these findings, our message is two-fold. Firstly, despite that the estimated MPCs among individual car purchasers is large, the implied aggregate MPCs are small once the appropriate conversion has been applied. We speculate that the difference may reflect different identification strategies and that our estimate reflects normal economic times. Secondly, the central role of a household's combined LTV ratio highlights that the workings of the credit and mortgage markets determine how fluctuations in house prices transmits to the macroeconomy not only in boom-bust episodes but also in normal times.

Bibliography

Adelino, M., Schoar, A., and Severino, F. (2015). Loan originations and defaults in the mortgage crisis: Further evidence. mimeo.

Adelino, M., Schoar, A., and Severino, F. (2016). Loan originations and defaults in the mortgage crisis: The role of the middle class. *Review of Financial Studies*, 29(7):1635–1670.

Aladangady, A. (2017). Housing wealth and consumption: Evidence from geographically linked microdata. *American Economic Review*, 107(11):3415–3446.

Angrist, J. and Krueger, A. B. (1992). The effect of age at school entry on educational attainment: An application of instrumental variables with moments from two samples. *Journal of the American Statistical Association*, 87:328–336.

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

Attanasio, O. P., Leicester, A., and Wakefield, M. (2011). Do house prices drive consumption growth? the coincident cycles of house prices and consumption in the uk. *Journal of European Economic Association*, 9(3):399–435.

Benmelech, Efraim, E., Meisenzahl, R. R., and Ramcharan, R. (2017). The real effects of liquidity during the financial crisis: Evidence from automobiles. *Quarterly Journal of Economics*, 132(1):317–365.

Berger, D., Guerrieri, V., Lorenzoni, G., and Vavra, J. (2017). House prices and consumer spending. *Review of Economic Studies, forthcoming*.

Berger, D. and Vavra, J. (2015). Consumption dynamics during recessions. *Econometrica*, 83(1):101–154.

Browning, M., Gørtz, M., and Leth-Petersen, S. (2013). Housing wealth and consumption: A micro panel study. *The Economic Journal*, 123:401–428.

Calomiris, C., Longhofer, S. D., and Miles, W. (2012). The housing wealth effect: The crucial roles of demographics, wealth distribution and wealth shares. NBER Working Paper number 17740.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747.

Campbell, J. Y. and Cocco, J. F. (2007). How do house prices affect consumption? evidence from micro data. *Journal of Monetary Economics*, 54:591–621.

Carroll, C. D. (2004). Housing wealth and consumption expenditure. Working paper.

Carroll, C. D., Otsuka, M., and Slacalek, J. (2010). How large are housing and financial wealth effects? a new approach. ECB working paper no. 1283.

Carroll, C. D., Otsuka, M., and Slacalek, J. (2011). How large are housing and financial wealth effects? a new approach. *Journal of Money, Credit and Banking*, 43(1):55–79.

Case, K. E., Quigley, J. M., and Shiller, R. J. (2005). Comparing wealth effects: The stock market versus the housing market. *Advances in Macroeconomics*, 5(1):1–32.

Chay, K. Y. and Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, 113(2):376–424.

Cooper, D. (2013). House price fluctuations: The role of housing wealth as borrowing collateral. *Review of Economics and Statistics*, 95(4):1183–1197.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.

Davidoff, T. (2013). Supply elasticity and the housing cycle of the 2000s. *Real Estate Economics*, 41(4):793–813.

Davidoff, T. (2016). Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors. *Critical Finance Review*, 5(2):177–206.

DeFusco, A. (2018). Homeowner borrowing and housing collateral: New evidence from expiring price controls. *Journal of Finance*, 73(2):523–573.

Disney, R., Gathergood, J., and Henley, A. (2010). House price shocks, negative equity, and household consumption in the united kingdom. *Journal of European Economic Association*, 8(6):1179–1207.

Fredriksson, P. and Öckert, B. (2013). Live-cycle effects of age at school start. *The Economic Journal*, 124:977–1004.

Gorea, D. and Midrigan, V. (2017). Liquidity constraints in the u.s. housing market.

Iacoviello, M. (2011). Housing wealth and consumption. Board of Governors of the Federal Reserve System, International Finance Discussion Papers no. 1027.

Inoue, A. and Solon, G. (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics*, 92(3):557–561.

Kaplan, G., Mitman, K., and Violante, G. (2016). Non-durable consumption and housing net worth in the great recession: Evidence from easily accesible data. NBER Working Paper number 22232.

Kaplan, G., Mitman, K., and Violante, G. (2017). The housing boom and bust: Model meets evidence. Working Paper.

Kommunikationsdepartementet (1996). Förutsättningar för avveckling av bromma flygplats. Regeringskansliets Offsetcentral, DS 1996:12.

Laerkholm Jensen, T. and Johannesen, N. (2017). Consumption effects of the 2007-2008 financial crisis: Evidence from households in denmark. *American Economic Review*, 107(11):3386–3414.

Luftfartsverket (2000). Flygplatskapaciteten i stockholmsregionen: sammandrag av lfvs slutredovisning januari 2000. LFV, Sweden.

Mian, A., Rao, K., and Sufi, A. (2013). Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics*, 128(4):1687–1726.

Mian, A. and Sufi, A. (2014). What explains the 2007-2009 drop in employment? *Econometrica*, 82(6):2197–2223.

Miljödomstolen (2006). Akt m1414-07, aktbilaga 1107, bilaga 7. Miljödomstolen, Nacka tingsrätt.

Miljööverdomstolen (2010). Luftfartsverket and naturvårdsverket v. länsstyrelsen i stockholms län. Svea Court of Appeal in Stockholm, case 1441-09.

Murphy, K. M. and Topel, R. H. (2002). Estimation and inference in two-step econometric models. *Journal of Business & Economic Statistics*, 20(1):88–97.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics*, 125(3):1253–1296.

Stockholm (1989). Bromma flygfält – en utredning om brommafältets framtid. Stockholms stad.

Stockholm (2008). Avtal med luftfartsverket om förlängning av upplåtelsen av stockholm-bromma flygplatsområde. Stockholms stad, available from http://insynsverige.se/documentHandler.ashx?did=43049 [Accessed: June 3, 2014].

Vestman, R. (2017). Limited stock market participation among renters and home owners. Working Paper.

Appendix A - Sample Restrictions in the Household Data Set

Table 4.9: Imposed restrictions on the household data set

Sample restrictions	# Households
1. Households that are coded as living in Stockholm municipality and own their main property in Stockholm in the end of 2006	41,147
2. Drop households if at least one household member owns more than one property in Stockholm	39,519
3. Drop households that own only untaxed property*	37,917
4. Drop households that own more than one property in Stockholm	37,898
5. Drop households that whose residential address is in an area of Stockholm that has no single-family houses	36,580
6. Drop households with inconsistent geocoding**	36,075
7. Drop households that own less than 50% of a property	35,397
8. Drop extreme outliers in CLTV***	35,386

Note: The table presents the number of households that remain after imposing consecutive sample restrictions.
*) Examples of untaxed property classes are boat houses, churches, and small allotments for private farming.
**) Our initial sample of 41,147 households have Stockholm municipality as primary residence. In some cases the geocoding is inconsistent with this.
***) The cut-off in terms of combined loan-to-value ratio (CLTV) is 1,000. Approximately 0.6% of the households have a CLTV exceeding 2.0. 4.4% of the households have a CLTV exceeding 1.0. The CLTV includes student loans.

Appendix B - Details on the real estate transactions data set

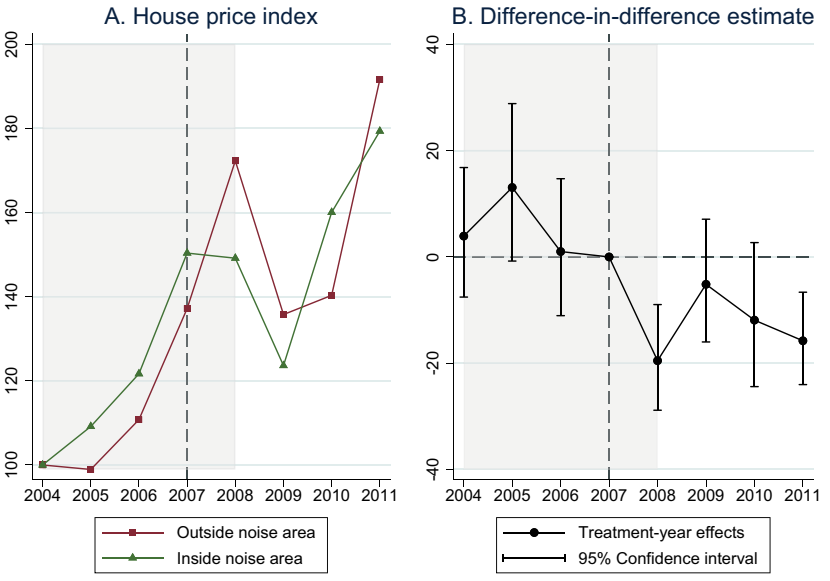
Residualization of Data

The data set on real estate transactions include approximately 2,500 transactions per year within the municipality of Stockholm. Statistics Sweden's official index FASTPI includes fewer transactions. The FASTPI index excludes properties with site-leasehold rights (rather than title deeds), transactions based on low prices, and transactions that display irregular discrepancies between tax value and transaction price. In our preferred specification we choose to strike a balance between maintaining representativeness of Stockholm's home owners and comparability with FASTPI by controlling for these factors in our preferred specification in Column (3) in 4.4. This is comparable to residualizing our data, rather than excluding observations all together. Hence, in order to compare our house price index to official house price indices we first residualize the transaction price against the following covariates: tax value, a dummy variable that indicates site-leasehold right, and a dummy variable for transaction price below 250 kSEK. To these residuals we add the average price. Figure 4.6 displays the equivalent of Figure 4.2 in the main text without prior residualization.

Comparison to official house price indices

We compare our house price index of Panel A in Figure 4.2 to official providers' indices. Figure 4.7 displays our house price index ("Outside and inside") next to a version that imposes similar sample restrictions as Statistics Sweden (SCB) does when constructing

Figure 4.6: Effect on house prices (non-residualized data)



Note: The figure displays the equivalent of Figure 4.2 in the main text without prior residualization.

their index FASTPI.[37] We drop all transactions with a price below 250 kSEK and all transactions that involve site-leasehold rights. We then residualize the data with respect to tax value, only. Next to these two versions of our index, the figure reports SCB's index and Valueguard's index HOX Stockholm Villa. The indexes are similar up until 2009. The discrepancy is notably larger in the years 2009 and onwards. One reason may be that average characteristics of transacted properties shifts due to the financial crisis which makes index construction more difficult. See Section 4.6 below. This is one reason to exclude 2008Q4 and later. Among all indices, SCB's is the least volatile one because of the heavy restrictions imposed on the input data.

No Composition Effects of Contract Renewal

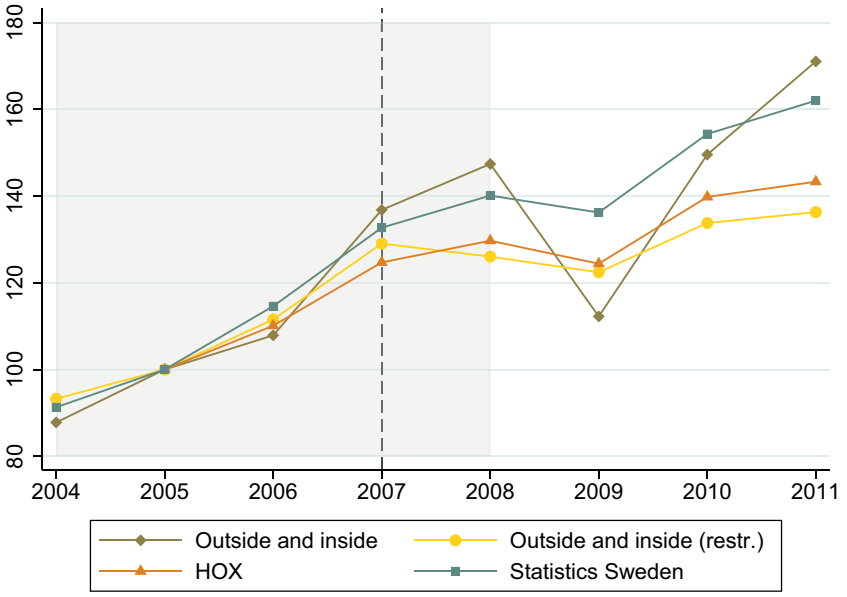
Figure 4.8 report DID estimates of the effect on properties' characteristics inside versus outside the noise area. The estimates indicate that the renewal of Bromma Airport's operating contract did not lead to compositional effects inside versus outside the noise area.

Fluctuations in House Characteristics

Figure 4.9 show average characteristics of transacted properties (Panels A to E) and transaction volume (Panel F). There are both temporary and longer lasting effects of the financial crisis.

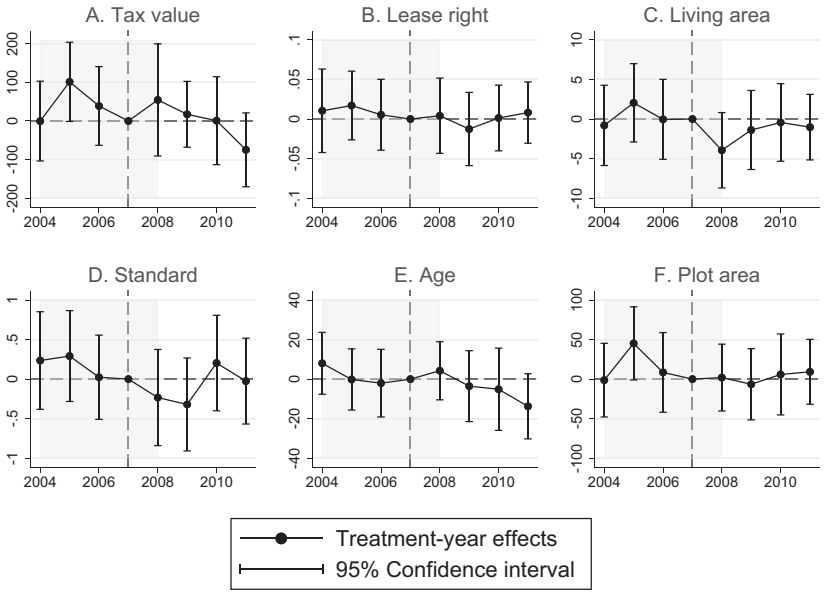
[37] We are grateful to Martin Verhage at Statistics Sweden for applying Statistics Sweden's method to the transactions in the municipality of Stockholm.

Figure 4.7: Comparison to official house price indices



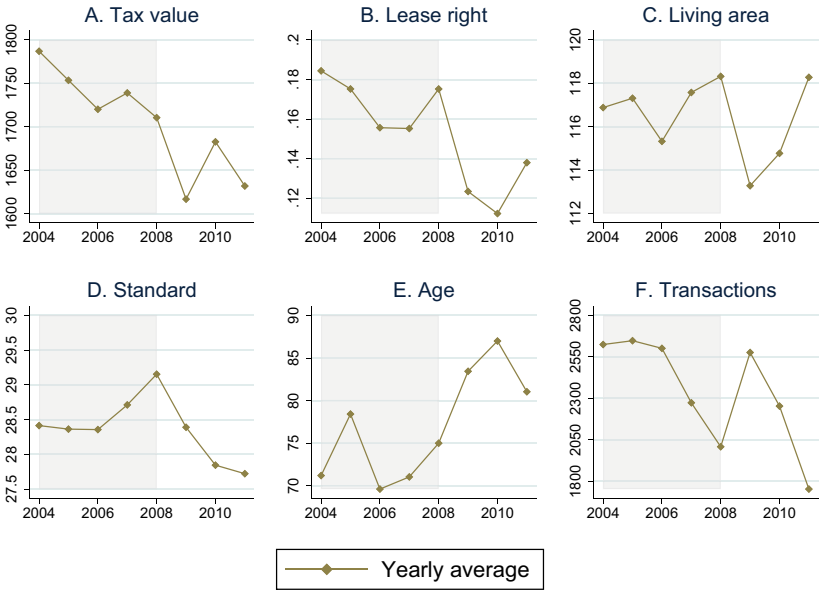
Note: The figure displays the house price index of Figure 4.2 ("Outside and inside") next to a version based on a restricted data set and next to Valueguard's index HOX Villa Stockholm Statistics Sweden's index FASTPI for the municipality of Stockholm. For all indices the timing is calendar year, as opposed to the timing in the main text.

Figure 4.8: Composition effects of contract renewal



Note: The figure displays DID estimates analogous to Figure 4.2 for six property characteristics.

Figure 4.9: Fluctuations in average characteristics



Note: Panel A to E in the the figure display average characteristics of transacted properties. Panel F report transaction volume.

Appendix C - Alternative Control Groups

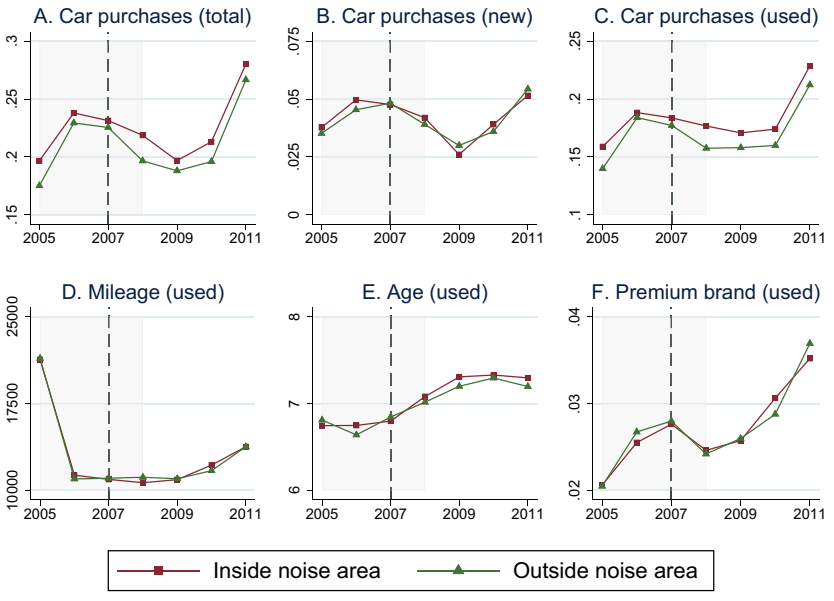
Table 4.10: Alternative control groups

	(1)	(2)	(3)	(4)	(5)
Noise area$_i$ × Post$_t$	-0.198***	-0.143***	-0.054	-0.085*	-0.136***
	(0.033)	(0.037)	(0.046)	(0.033)	(0.034)
Noise area$_i$	0.087*	0.063	0.058	0.066	0.064
	(0.045)	(0.045)	(0.046)	(0.047)	(0.044)
Observations	10,599	8,210	5,897	5,471	9,547
R-squared	0.338	0.371	0.356	0.357	0.335
Control	South-E.	South-W.	North-E.	North-W.	<5,000m
Avg. distance to NC in control	4,345m	3,438m	1,878m	1,858m	3,328m
2004Q1-2008Q3	Yes	Yes	Yes	Yes	Yes
Restricted	Yes	Yes	Yes	Yes	Yes
Parish FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Note: Estimates based on the full treatment group but different geographical control groups. Clustering of standard errors at the level of the 250 × 250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix D - Details on Car Transactions

Figure 4.10: Car purchases: extensive margin and substitution effects



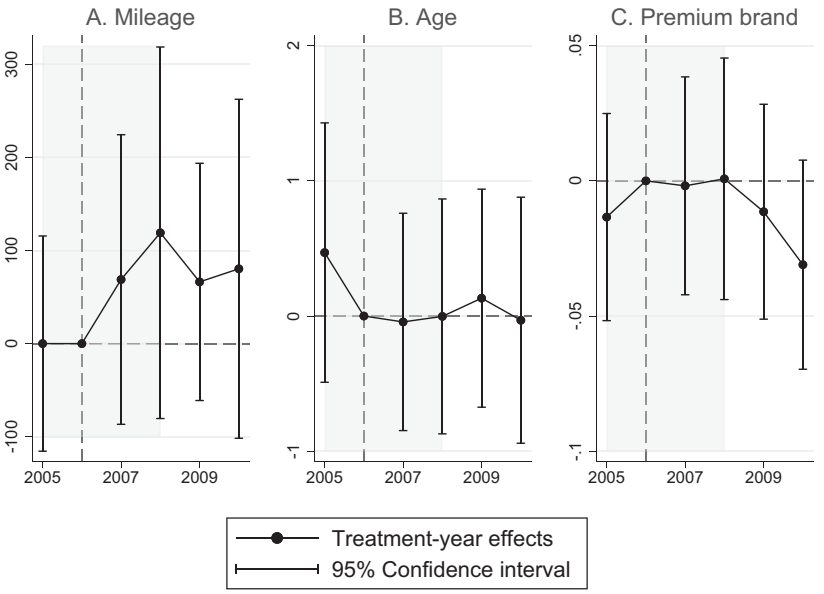
Note: The figure shows 12-months averages inside and outside the noise area, corresponding to the DID estimates in Figure 4.4. The timing on the horizontal axis is shifted one quarter. That is, 2007 refers to 2006Q4–2007Q3 and so forth. Premium brands refer to Audi, BMW, and Mercedes. The shaded grey area indices the main sample period (2004Q1-2008Q3).

Table 4.11: Effects on cars – extensive margin and used cars

	(1)	(2)	(3)	(4)	(5)	(6)
	New & Used	New	Used	Mileage	Age	Used premium
Noise area$_i$ × Post$_t$	0.002	0.000	0.002	-594.767	0.001	0.000
	(0.002)	(0.001)	(0.002)	(520.838)	(0.194)	(0.001)
Noise area$_i$	0.002	0.000	0.000	62.429	0.214	-0.000
	(0.002)	(0.001)	(0.002)	(653.967)	(0.147)	(0.001)
N	530,790	530,790	530,790	21,927	21,972	530,790
R-squared	0.015	0.003	0.014	0.061	0.029	0.004
2005-2008Q3	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports DID estimates corresponding to Figure 4.4. The set of control variables includes age of household head, household size, financial wealth, housing wealth, total debt, labor income, and capital income. Clustering of standard errors at the level of the 250 × 250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4.11: Car sales: extensive margin and substitution effects



Note: The figure shows 12-months averages inside and outside the noise area, analogous to the DID estimates in Figure 4.4 for households' sales. The timing on the horizontal axis is shifted one quarter. That is, 2007 refers to 2006Q4–2007Q3 and so forth. Premium brands refer to Audi, BMW, and Mercedes. The shaded grey area indices the main sample period (2004Q1-2008Q3).

Table 4.12: Effect on the price of new cars: dose response

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Post}_t \times 0$	-0.087**						
	(0.046)						
$\text{Post}_t \times (0 \quad 500]$	-0.048						
	(0.043)						
$\text{Post}_t \times (500 \quad 1000]$	-0.037						
	(0.048)						
$\text{Post}_t \times \log^{10}(\text{dist}^2)$		0.035**					
		(0.016)					
$\text{Post}_t \times 0$			-0.028				
			(0.023)				
$\text{Post}_t \times [0 \quad 200]$				-0.058*			
				(0.033)			
$\text{Post}_t \times [0 \quad 400]$					-0.044		
					(0.034)		
$\text{Post}_t \times [0 \quad 600]$						-0.074**	
						(0.033)	
$\text{Post}_t \times [0 \quad 800]$							-0.072***
							(0.031)
N	4,794	4,794	4,794	4,794	4,794	4,794	4,794
R-squared	0.078	0.077	0.077	0.078	0.078	0.078	0.079
2005Q1-2008Q3	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

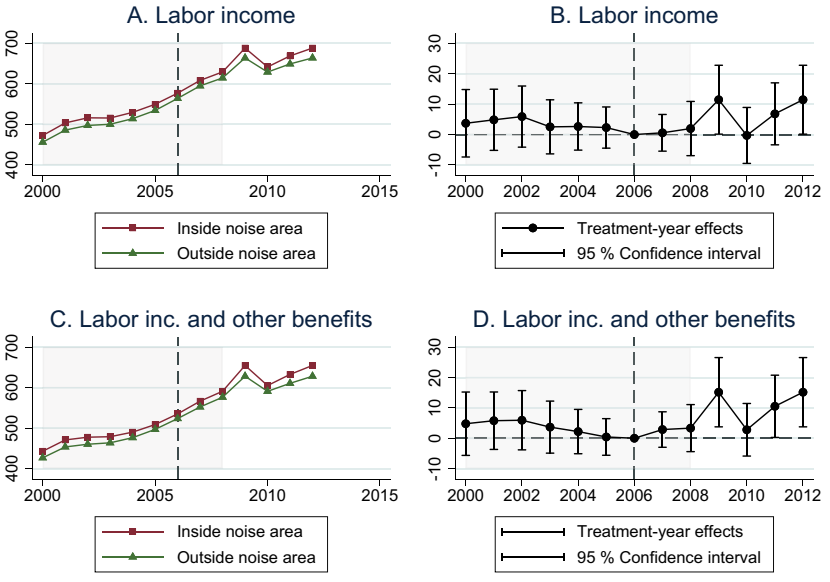
Note: The set of control variables includes age of household head, house hold size, financial wealth, housing wealth, total debt, labor income, and capital income. Clustering of standard errors at the level of the 250×250 meter grid. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix E - Additional Statistics

Figure 4.12: Geographical distributions relative to the noise contour

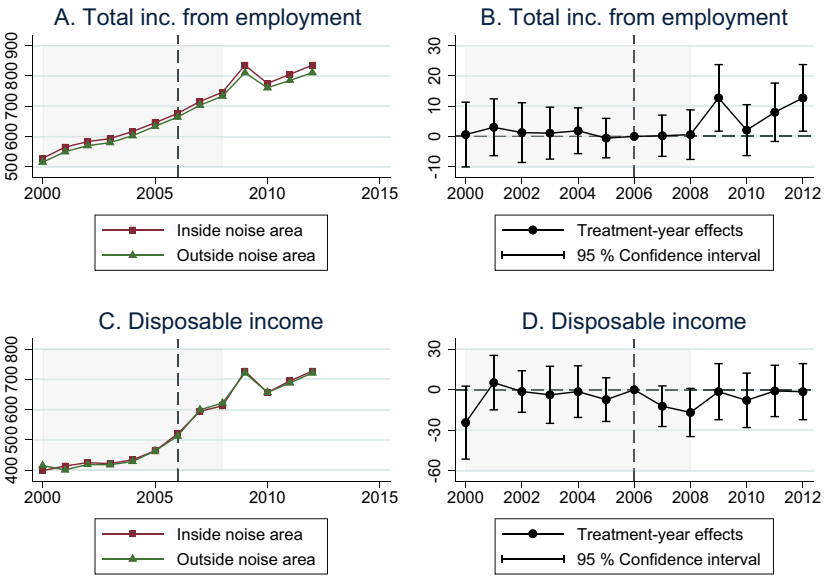
Note: The figure displays the geographical distributions of house transactions, households, and car transactions.

Figure 4.13: Labor income and other employee benefits



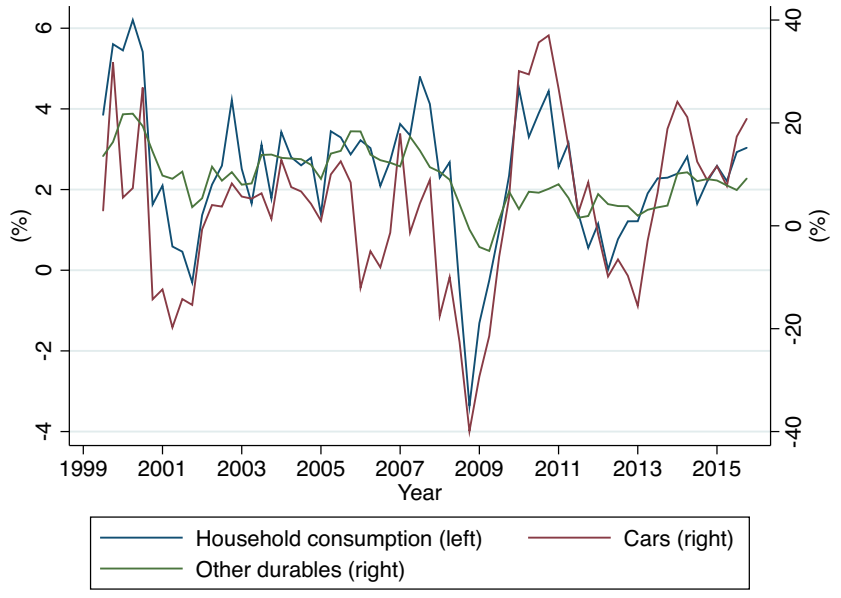
Note: The figure displays the evolution of labor income (Panels A and B) and labor income plus other employee benefits (Panels C and D).

Figure 4.14: Labor income and other employee benefits



Note: The figure displays the evolution of total income from employment (Panels A and B) and disposable income, defined as gross labor income plus capital income plus transfers minus taxes (Panels C and D).

Figure 4.15: Household consumption growth (Sweden)



Note: The figure displays year-on-year consumption growth for Swedish households.

Appendix F - Previous Estimates

Table 4.13: Estimates of housing wealth effects from previous literature

Study	Country	Elasticity	MPC	MPC cars
Cambell and Cocco (2007)*	UK	0-1.7	–	–
Disney et al (2010)	UK	–	0.01	–
Carrol (2011)**	US	–	0.02-0.09	–
Browning et al (2013)***	Denmark	0.13	0.05	–
Cooper (2013)	US	0.06	0.06	–
Mian et al (2013)****	US	0.13-0.25	0.054	0.023
Kaplan et al (2016)*****	US	0.24-0.36	–	–
Aladangady (2017)	US	–	0.047	–
Guerrieri and Iacoviello (2017)	US	–	0.24	–
DeFusco (2018)******	US	–	0.04-0.13	–

Note: The table presents estimates from previous studies. *) 1.7 is for older homeowners, 0 for young renters. Estimates need to be scaled with the proportion of the groups owning real estate for them to attain comparable aggregate estimates. **) Lower range of estimate is the direct effect, upper range is long run estimate. ***) Effects only for subsample (young and constraint). Estimates should be scaled with the proportion of such households to attain comparable aggregate estimates. ****) Scaled with H/NW, i.e. the housing wealth to net worth ratio. *****) Should be scaled with H/NW. ******) Assuming 100% consumption out of an increase in debt.

Sammanfattning (Swedish Summary)

Denna avhandling består av fyra självsående kapitel.

De första två kapitlen, som avhandlingen är döpt efter, handlar om skolval. Detta ämne har varit mitt huvudsakliga forskningsintresse de senaste tre åren utgör också grunden för min framtida forskningsagenda. Jag har, i båda dessa projekt, samarbetat med Elisabet Olme. Den övergripande motivationen för både dessa studier är den snabba och globala expansionen av skolvalsprogram vi sett under de senaste decennierna. Forskare inom olika fält har gjort stora framsteg när det kommer till att förstå konsekvenserna av denna utveckling. Vi har också kommit långt i arbetet att besvara hur denna typ av system optimalt bör utformas. Det finns dock fortfarande ett stort antal frågor där vi inte har tillräckliga svar. På grund av databegränsningar gäller detta särskilt frågor där det inte finns något entydigt teoretiskt svar. Det är också klart att praktiken laggar efter forskningen och jag tror att de flesta skolvalsforskare skulle hålla med mig i påståendet att alla skolvalsprogram som för närvarande används har brister. De två första kapitlen i denna avhandling kan betraktas som mitt första bidrag till den pågående insatsen för att försöka rätta till detta.

Det första kapitlet undersöker om information om skolkvalitet

kan påverka skolvalsbettende och elevallokering. Kapitlet bygger på ett stort fältexperiment som vi genomförde i Linköpings kommun år 2016. Vi randomiserade eleverna som skulle börja högstadiet in i en behandlingsgrupp och en kontrollgrupp. Hushållen i behandlingsgruppen fick information om de tillgängliga skolornas absoluta och, för elevkomposition, justerade resultat på de nationella proven. Vi finner ingen signifikant effekt av informationen på andelen elever som väljer bort från sin närmaste skola eller på kompositionen av dem som gör det. Vi finner dock att andelen hushåll som väljer en högpresterande skola ökar med 5,2 procentenheter (motsvarande 12 procent). Denna effekt drivs främst av hushåll med svensk bakgrund och högutbildade hushåll. Med hjälp av kommunens faktiska allokeringsmekanism simulerar vi hur denna förändring i valbeteende översätts till förändringar i elevallokering. Antalet elever som placeras på högpresterande skolor ökar men det förändrade valbeteendet slår inte igenom helt på grund av skolornas kapacitetsbegränsningar. ökningen drivs återigen av högutbildade hushåll och hushåll med svensk bakgrund. Dessa hushåll knuffar, i viss utsträckning, bort elever med utländsk bakgrund från de högpresterande skolorna. Sammantaget leder detta till en ökning av klyftan i skolprestation mellan socioekonomiskt starkare och svagare hushåll. Den inledande effekten på elevallokeringen propagerar också genom systemet. Eleverna som nu kommer in på de högpresterande skolorna lämnar efter sig lediga platser på översökta medelpresterande skolor. Dessa platser fylls, till stor del, av elever med utländsk bakgrund. Detta i kombination med det faktum att den genomsnittliga andelen av elever med utländsk bakgrund i de högpresterande skolorna är ganska hög, leder till en signifikant minskning i skolsegregationen i termer av migrationsbakgrund. Skolsegregationen i termer av föräldrars utbildningsbakgrund verkar dock öka något då studenter från högutbildade hushåll samlas i de högpresterande skolorna.

I det andra kapitlet utvärderar vi hur skolsegregationen och studentens välfärd påverkas när man justerar prioritetsstrukturerna i ett skolvalsystem. [38] Vi använder faktiska valdata från Botkyrka kommun för att estimera föräldrars preferenser för ett antal skolattribut. Dessa preferenser används simuleringar av kontrafaktiska allokeringar av studenter under olika typer av prioritetsstrukturer. Vi utvärderar tre olika typer av prioritetsstrukturer, en som bygger på närhet till skolan, en baserad på lotteri och ett baserat på så kallade mjuka kvoter eller reserverade platser. Resultaten tyder på att prioritetsstrukturerna påverkar graden av skolsegregation. När man reserverar platser för olika grupper blir skolorna mindre segregerade både vad gäller föräldrarnas utbildningsnivå och utländsk bakgrund jämfört med lotteri- och närhetsbaserade prioritetsstrukturer. Vi finner att kostnaderna i termer av genomsnittlig välfärd för denna minskade segregation är begränsad. Vidare finner vi att effekterna på välfärd av de olika prioritetsstrukturerna skiljer sig mellan olika grupper av hushåll. Högutbildade hushåll med svensk bakgrund gynnas mest av en närhetsbaserad prioritetstruktur medan lågutbildade hushåll med svensk bakgrund och högutbildade hushåll med utländsk bakgrund gynnas mest av en prioritetstruktur som bygger på reserverade platser.

De två sista kapitlen i denna avhandling är inte relaterade till skolval. De redovisar resultaten från två projekt som jag påbörjade innan jag valde att rikta in mig på utbildningsekonomi i allmänhet och skolval i synnerhet.

Det tredje kapitlet undersöker vad som händer om man informerar människor om könsskillnaderna i benägenheten att tävla och den ineffektivitet som denna könsskillnad ger upphov till. Ett antal experiment där deltagarna fick välja om de skulle tävla eller inte genom-

[38] Prioritetsstrukturerna är, enkelt uttryckt, reglerna som avgör vem som har företräde till översköta skolor.

fördes på en Nacka gymnasieskola. Ett slumpmässigt urval av deltagarna informerades om könsskillnaderna i benägenheten att tävla och de kostnader som denna skillnad medför. Bland deltagarna som inte fick någon information var män mycket mer benägna att tävla än kvinnor. Det fanns också betydande ineffektivitet som kom av att lågpresterande män valde att tävla och att högpresterande kvinnor valde att inte göra det. Bland deltagarna som fick informationen var kvinnor mer benägna att tävla än män och andelen felaktiga markant mindre. Minskningen i andelen felaktiga val drivs av högpresterande kvinnor och lågpresterande män. Bland lågpresterande kvinnor ökade andelen felaktiga val.

Det fjärde kapitlet är skrivet tillsammans med Roine Vestman och Björn Tyrefors Hinnerich och är ett försök att estimera bostadsförmögenhetseffekten. Vi utnyttjar ett kvasi-experiment som uppstod i Stockholm 2007 när kontraktet för Bromma flygplats oväntat förlängdes till 2037. Vi estimerar att huspriserna faller cirka 16 procent på fastigheter nära flygplatsen som en konsekvens av detta. Denna prisvariation är ideal för att identifiera bostadsförmögenhetseffekter eftersom den är lokal och oberoende av makroekonomiska förhållanden. Vi studerar hur dessa hushålls bilköp påverkas av detta prisfall och finner att det genomsnittliga priset för nya bilar köpta av dem vars hus fallit i värde minskar med cirka 7 procent. Korrigerar man för andelen nybilsköpare blir dock den totala effekten på konsumtion liten, ca 0.2 öre per krona som husen fallit i värde. Denna effekt är betydligt mindre jämfört med de flesta tidigare studier. Vi diskuterar ett antal anledningar till detta, inklusive det faktum att vårt estimat sannolikt är mycket närmare en sann partiell jämviktseffekt. Effekten är koncentrerad till husägare med en skuldsättningsgrad mellan 0.6 och 0.8, vilket indikerar att effekten kommer av lånerestriktioner.