The effects of value-added quality information on the outcomes of school choice – a further development and application of the micro-simulation approach

Dissertation

zur Erlangung des akademischen Grades des Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)

vorgelegt von Philipp Stütze

an der

Universität Konstanz

Sektion Politik – Recht – Wirtschaft

Fachbereich Wirtschaftswissenschaften

Tag der mündlichen Prüfung: 08.05.2018
1. Referent: Prof. Dr. Heinrich Ursprung
2. Referent: Prof. Dr. Guido Schwerdt
Prüfungsvorsitzende: Jun. Prof. Dr. Luna Bellani
Acknowledgments:

First and foremost, I would like to thank Prof. Dr. Ursprung for giving me plenty of rope while working on this thesis while being there to offer time and advice whenever I needed it. I also want to thank Mr. Ursprung for his advice to turn this topic into a full thesis, so as to make it worth the considerable effort of obtaining the data and devising, calibrating and running the simulations that provided the results. I would also like to thank Mr. Ursprung for the idea of going for a Ph.D. in the first place and for encouraging me that this undertaking is feasible, although my later wife lived in Paris at the time and for his attitude that “I do not care where or when you work as long as you make progress”. This allowed me to indeed maintain a long-distance relationship for years while making progress. Finally, I want to thank him for his unlimited patience while I struggled for almost 6 years to find the time to finish this thesis while working on a full time job and having small kids.

And I would like to thank Prof. Dr. Guido Schwerdt and Prof. Dr. Luna Bellani for accepting to be members of the examination committee on short notice.

I would also like to thank the Department of Economics at the University of Konstanz and the LGFG scholarship for providing me with the financial means to pursue this goal.

I owe thanks to Dr. Thomas Fuchs for a first draft of a micro-simulation on school choice. Although I doubt that a single line of his original code survived unaltered, his work gave me many ideas on how to program a micro-simulation using STATA.

Also, I would like to thank for valuable advice gained from comments during and after presentations of earlier stages of my research at the brown-bag seminars and the academic writing workshops in Constance and the University of Zurich, from Prof. Dr. Ludger Wößmann and the members of the department of Human Capital and Innovation at the Ifo Institute Munich and from conversations with members of the Chair of Political Economy and the Department of Economics in Constance and from the participants of the doctoral courses for education economics that I attended at the University of Zurich.

I would also like to thank my parents for their support and encouragement throughout my academic education.

Finally, I want to thank my wife for her support, encouragement, love and patience during 5 years of a long-distance relationship that encompassed our wedding and most of a pregnancy. And I thank my wife and children for allowing me to spend more than half my annual leave entitlement for an additional 6 consecutive years, as well as countless hours at evenings and weekends, for finishing this thesis while already working on a full time job.
List of figures ........................................................................................................ viii
List of tables ........................................................................................................... x
Abbreviations ........................................................................................................... xi
1. Introduction ......................................................................................................... 1
2. Friedman’s mechanism in more detail .............................................................. 1
3. Friedman’s mechanism vs. the real world ......................................................... 3
  3.1. Parental school choice behaviour ................................................................. 3
    3.1.1. Hypotheses ............................................................................................... 3
    3.1.2. Empirical evidence .................................................................................... 4
    3.1.3. Summary on school choice behaviour of parents ....................................... 9
  3.2. Behaviour of teachers and schools ............................................................... 9
    3.2.1. Hypotheses ............................................................................................... 9
    3.2.2. Empirical evidence .................................................................................... 10
    3.2.3. Summary concerning the role of schools and teachers in school choice ....... 18
  3.3. The effects of institutions and regulations on school choice systems .......... 19
    3.3.1. Hypotheses ............................................................................................... 19
    3.3.2. Empirical evidence .................................................................................... 22
    3.3.3. Summary on the effects of institutions and regulations ............................. 30
  3.4. Students and effects of changing schools ................................................. 31
    3.4.1. Hypotheses ............................................................................................... 31
    3.4.2. Empirical evidence .................................................................................... 32
    3.4.3. Summary on students and the effects of changing schools ....................... 40
  3.5. Information .................................................................................................... 41
    3.5.1. Different possibilities of measuring and presenting academic school quality 41
    3.5.2. Observations and empirical evidence on academic quality measures ......... 46
    3.5.3. Summary on information ......................................................................... 51
4. Approaches to analyzing the effects of school choice in previous literature ......... 53
  4.1. General difficulties ....................................................................................... 53
    4.1.1. Differences in institutional settings and characteristics .......................... 53
    4.1.2. Interdependence of decisions .................................................................... 54
    4.1.3. Heterogeneous actors ............................................................................... 55
  4.2. Empirical approaches .................................................................................... 56
8.1.3. Computation of the educational productivity and comparison to average outcomes

8.2. Differentiation of students into types

8.3. Estimation of the educational production function

8.4. Estimating the parental decision function for: which students choose schools?
  8.4.1. Explanatory variables for parental decision function 1
  8.4.2. Estimation of parental decision function 1
  8.4.3. Cut-off values for the predictions based on the parental decision function 1

8.5. Estimating the parental decision function for: which schools are chosen?
  8.5.1. Explanatory variables for parental decision function 2
  8.5.2. Estimation of parental decision function 2
  8.5.3. Predictions for use in the simulation based on parental decision function 2

8.6. Modelling the impact of distance

8.7. Competitive pressure
  8.7.1. The impact of relative academic quality on competitive pressure
  8.7.2. The impact of distance on competitive pressure
  8.7.3. The measure for competitive pressure

8.8. The productivity determination function
  8.8.1. Estimation of the productivity determination function
  8.8.2. Limit yearly changes in educational productivity
  8.8.3. Effects of a directly observable educational productivity
  8.8.4. Adjusting the productivity determination function to the case of observable productivity

8.9. Imputation of dropout status for the cohort of 2005

8.10. The estimated model

8.11. Robustness measures
  8.11.1. Randomization of the EPF
  8.11.2. Choice functions

9. The simulation
  9.1. Simulation approach
  9.2. The two scenarios
    9.2.1. Scenario one: passive schools and pure sorting effects
List of figures

Figure 8-1: Distribution of educational productivity across schools, weighted by number of students 142
Figure 8-2: Scatter-plot of average test results and educational productivity of schools 144
Figure 8-3: Distribution of the share of low-income students across schools 160
Figure 8-4: Distance to the attended school, dataset 164
Figure 8-5: Distance to the attended school, simulated 165
Figure 8-6: Distribution of apparent quality gain from school-change 167
Figure 8-7: Weight of competing schools based on distance, observed values and approximated function 170
Figure 8-8: Effect of competitive pressure on educational productivity 172
Figure 8-9: Effect of free capacity on educational productivity 173
Figure 8-10: Distribution of year on year changes in observed educational productivity 175
Figure 8-11: Effect of competitive pressure on educational productivity 183
Figure 9-1: Scenario 1: Distribution of average gains in productivity as compared to a scenario without choice 198
Figure 9-2: S1: Average gain in productivity for school changers 199
Figure 9-3: S1: Average productivity gains over time 200
Figure 9-4: S1: Distribution of average gains in test scores as compared to no-choice scenario 202
Figure 9-5: S1: Score gains by year 203
Figure 9-6: S1: Score gains, distribution for choosers 204
Figure 9-7: S1: Score gains, distribution for non-choosers 204
Figure 9-8: S1: Gains in average peer ability, distribution for choosers 206
Figure 9-9: S1: Gains in average peer ability, distribution for non-choosers 206
Figure 9-10: S1: Distribution of share of school-changers 207
Figure 9-11: S1: Distribution of innate ability of school-changers across simulation outcomes 208
Figure 9-12: S1: S1: Distribution of sorting by innate ability for high ability students 210
Figure 9-13: S1: S1: Distribution of sorting by innate ability for low ability students 211
Figure 9-14: S1: Distribution of sorting by race for African Americans and Hispanics 212
Figure 9-15: S1: Distribution of sorting by race for White students and by income status 212
Figure 9-16: S1: Average share of free places at more productive schools by turn, based on expected and experienced level 213
Figure 9-17: S1: Average share of free places at high-scoring schools by turn, based on expected and experienced level 214
Figure 9-18: Scenario 2: Distribution of average gains in productivity as compared to a scenario without choice 217
Figure 9-19: S2: Distribution of average gains in educational productivity for choosers 218
Figure 9-20: S2: Average gain in productivity by year as compared to scenario without choice 219
Figure 9-21: S2: Distribution of average gains in test scores as compared to no-choice scenario 221
Figure 9-22: S2: Score gains by year 222
Figure 9-23: S2: Distribution of score gains for choosers 223
Figure 9-24: S2: Distribution of score gains for non-choosers 223
Figure 9-25: S2: Distribution of gains in peer ability for choosers 225
Figure 9-26: S2: Distribution of gains in peer ability for non-choosers 226
Figure 9-27: S2: Distribution of share of school-changers 227
Figure 9-28: S2: Distribution of innate ability of school-changers across simulation outcomes 229
Figure 9-29: S2: Distribution of sorting for high innate ability 230
Figure 9-30: S2: Distribution of sorting for low innate ability 230
Figure 9-31: S2: Distribution of sorting by race for African Americans and Hispanics 231
Figure 9-32: S2: Distribution of sorting by race for White students and by income status 232
Figure 9-33: S2: Average share of free places at more productive schools by turn, based on expected and experienced level 233
Figure 9-34: S2: Average share of free places at high-scoring schools by turn, based on expected and experienced level 234
**List of tables**

Table 7-1: Average characteristics of students in the dataset for preparatory estimations 130  
Table 7-2: Average characteristics of students in the dataset for the simulations 131  
Table 8-1: Estimation of the auxiliary educational production function 137  
Table 8-2: Auxiliary educational production function by subject 138  
Table 8-3: Auxiliary educational productivity function by ability level 146  
Table 8-4: Educational production function 147  
Table 8-5: Educational production function by race 149  
Table 8-6: Parental decision function 1: Choice or not? 153  
Table 8-7: Parental decision function 1 by race 154  
Table 8-8: Parental decision function 2: Which school? 158  
Table 8-9: Parental decision function 2 by race 161  
Table 8-10: Estimation of the productivity determination function 171  
Table 8-11: Productivity determination function including predicted scores of student intake 174  
Table 8-12: Average characteristics of predicted dropouts 185  
Table 8-13: Average characteristics of dropouts who are identified in the dataset 185  
Table 8-14: Logit estimation of dropout status 186
Abbreviations

ACT: American College Test
CGE: Computable General Equilibrium
CPS: Chicago Public Schools
EPF: Educational Production Function
IEP: Individual Education Plan
ISAT: Illinois Standards Achievement Test
Meetexceed: the share of students who met or exceeded state standards in the PSAE
NCLB: No Child Left Behind Policy
NCTT: No Choice This Turn, used to identify the results of the individual choice decision of one student, holding the choices of all other actors constant
OUT-case: case in the simulation where quality measures are based on the outcome measure of the share of students who met or exceeded state standards
PROD-case: case in the simulation where quality measures are based on educational productivity
PSAE: Prairie State Achievement Examination
REA: Office of Research, Evaluation, and Accountability of the CPS
SPF: School Productivity Function
1. Introduction

“Could school choice be a tide that lifts all boats?”
(Hoxby 2002a, title)

“While information may be the Achilles’ heel of choice, many of the existing treatments of the flow of information have been limited.”
(Teske/Schneider 2001, p.624)

“Many accountability systems that appear reasonable at first glance perform in perverse ways when test score measures are imprecise.”
(Kane/Staiger 2002, p.91)

The three citations above set up the starting point and motivation for this dissertation. I will refer to each in turn and by doing so describe the aim and approach of this dissertation.

One of the most important arguments that proponents of school choice put forward is that the right to choose schools will increase the average educational productivity of schools. Through this increase in productivity, school choice is supposed to “...lift all boats...” (Hoxby 2002a, title) by benefiting all students, whether they actively choose schools or not. This increase in educational productivity was presented for example in 1955 by Milton Friedman (Friedman 1955) as one beneficial outcome of increased choice, originally through state-sponsored vouchers for education. The mechanism behind this increase in productivity is the underlying assumption of much research on school choice and was summarized for example by Caroline Hoxby (Hoxby 2003b, p.17f).

The mechanism works as follows: If students are given the right to freely choose schools, and money follows students, this creates competition in a quasi-market for education. To survive in this market, schools have to attract students. As parents want their children to learn as much as possible for a given effort, they will prefer, ceteris paribus, those schools with a higher educational productivity. Less productive schools have to improve productivity or they shrink and are replaced by more productive ones. The parental preference for productivity combined with the quasi-market thus creates competitive pressure on all schools that results in an overall increase in school productivity.
The empirical evidence on the effects of school choice is however unclear and partially conflicting. The main reason for diverging empirical results on the effects of school choice is, that there are several implicit assumptions behind the mechanism described above, which I will call Friedman’s mechanism throughout this dissertation\(^1\). And these assumptions are not necessarily met in a given school choice system. I will identify these implicit assumptions in chapter 2. In chapter 3 I will survey the evidence in the literature on school choice, to analyze how frequently and to what extent these assumptions are met in real world school choice systems. I will show that several of these assumptions are not met in many existing school choice systems. In particular, one crucial assumption is often not met: the ability of parents to identify the most productive schools. If parents cannot identify the most productive schools, then active school choice does not necessarily lead to academic gains for individual students and schools do not necessarily benefit much from having a high productivity. Thus, information about the educational productivity of schools is the “…Achilles’ heel of choice…” as cited above (Teske/Schneider 2001).

The remainder of this dissertation (starting with chapter 4) aims at identifying how exactly the outcomes of choice are affected if parents cannot observe the productivity of schools and how the effects of choice would change, if the educational productivity of schools would be made observable to parents.

However, school choice has some peculiarities that make it a hard subject to study. Among these peculiarities is the fact, that there is a strong interdependency between the decisions of most actors. The second of these peculiarities of school choice is, that there are many institutional settings and external conditions by which an individual school choice system might differ from the next one. Several of these settings and external conditions can single-handedly break the mechanism that was first described by Friedman. If set in the wrong way, some of these settings might even induce the school choice system to produce unfavourable results. One example for such a setting is the wrong type of information about the quality of schools, that is the focus of the simulations at the end of this dissertation (in chapter 9) and that can induce an otherwise seemingly reasonable school choice system to “…perform in perverse ways...” as cited above (Kane/Staiger 2002, p.91). The analysis of school choice is further complicated by cross-effects between individual settings and external conditions.

---

\(^1\) There are more compact and clearer descriptions of this mechanism by other authors, and Friedman does not fully and explicitly describe it in his 1955 text (Friedman 1955) that proposes only choice through vouchers. But the missing parts of the mechanism can be inferred from the remainder of his text and the conclusions that he draws. Moreover, Friedman is, to my knowledge, the first economist to describe this mechanism in an official publication. Therefore, I decided to attribute the mechanism to him.
Chapter 4 provides an in-depth analysis of these peculiarities and how they combine to render conventional theoretical and empirical approaches almost useless for the analysis of the impact of different types of information on the outcomes of school choice. Consequently, I need to use another approach that is capable to deal with the peculiarities of school choice. Computational micro-simulation models are such an approach. Chapter 5 presents this approach and three models that were previously employed to analyze school choice using a micro-simulation approach. These models present a great step forward but are lacking detail and are not based strongly enough on empirical findings to yield the insights that I intend to gain with this dissertation. I therefore develop my own micro-simulation model in chapter 6, that is loosely based on an approach by Thomas Fuchs (Fuchs 2007). Chapter 7 presents a dataset from the Chicago Public Schools (CPS) that I will use to estimate and calibrate the model. This estimation and calibration of the model are done in chapter 8. In chapter 9 I apply the simulation model to the analysis of the effects of making educational productivity of individual schools observable and present the results. Chapter 10 concludes and gives an outlook on how the refined micro-simulation approach, which I have developed in this dissertation, could be used to gain further insights into the effects of school choice. This chapter also gives an outlook on how the effects of school choice are affected by changes in individual institutional settings and external conditions.

Before I start with the main part of the dissertation, I will define two concepts, declare how I will handle the gender of groups of persons that I will mention recurrently and give information regarding the software that I used and how to access and replicate my results.

Firstly, I define a kind of school productivity, which I will be using throughout this dissertation. The educational productivity of schools is the contribution that a school makes to the educational attainment of a student for a given level of financing, for given characteristics of an individual student and for given aggregate characteristics of her peers. This differs from another common use of school productivity, which is defined as the student outcome per dollar of education financing. This monetary productivity does not account for characteristics in the student intake and is thus strongly affected by systematic differences in these characteristics. A definition of educational productivity that represents the value-added that schools provide to the educational attainment of students, comes closer to what Milton Friedman describes, than a measure that increases only because the student intake of a school comes from a wealthier neighbourhood.
Secondly, I define **value-added**: This term has been defined twice in the education economics literature. I use the term value-added, as in value-added (school) quality measures, to describe the contribution of the educational quality of a school to the academic achievement of its students. Value-added has also been used in the education economics literature to define the gains in test scores that a student achieved within a defined time period, usually one year.

Thirdly, I state how I will handle the **gender** of persons that I mention in the dissertation: To avoid clumsy expressions such as his/her and he/she, I will always use the same gender for the same type of actor. Students are always female, teachers always male. All other actors, like parents and school administrators, are never mentioned as individuals but as groups, so that I do not have to decide on a gender.

Finally, some information about the software that I used and replication of results: I used the software **STATA** (Stata 10) for estimations and the simulation. I strove to make my results replicable by including the datasets and command files that I used and extensive comments in these command files. For more information see appendix 1.
2. Friedman’s mechanism in more detail

In this chapter I analyze the mechanism that was presented by Milton Friedman in more detail. The argument is as follows: Give parents the right to choose the schools for their children. This creates a quasi-market for education. To survive in this market, schools have to attract students. As parents want their children to learn as much as possible for a given effort, they will prefer those schools with a high productivity. The parental preference for productivity combined with the quasi-market thus creates competitive pressure on all schools to raise productivity. Thus school choice leads to an overall increase in school productivity.

This argument seems straightforward. But the described mechanism is based on assumptions, many of which are not explicit on the first glance. To identify these assumptions, I will have a closer look at the actors (parents, schools/teachers and students) and at the role of regulation and information in school choice systems. In this chapter, I will list the assumptions that are implicit to the mechanism as presented by Milton Friedman. In chapter 3 I will describe these assumptions in more detail and break them down into identifiable items. Then I will present evidence from empirical literature on how often and how completely the identified assumptions are fulfilled in existing school choice systems.

I start by sketching implicit assumptions about parents and their behaviour.

For school choice to lead to an increase in school productivity:

Parents first need to be able to identify schools that have a high productivity. Then, sufficiently many parents have to choose these schools. And those students who switch schools have to be of a kind that schools want to attract.

Students have to be willing to attend another school than the closest one. At least, they should not be able to resist a school choice that their parents want them to make. As a consequence of the assumed parental preference for schools with a high productivity, students should benefit academically from active school choice.

Teachers and schools have to feel enough pressure so that they want to change their behaviour in a way that increases productivity. For such an increase to realize, several conditions have to be met. To start with, teachers have to know how to increase productivity. Secondly, teachers have to expect that they are able to increase productivity enough to ease competitive pressure. Thirdly, teachers have to expect that the benefits of a productivity increase will be worth the required effort. Finally, it has to be more rewarding for teachers to
increase productivity than to ease competitive pressure by other means that require less effort. One such alternative mean could be cheating.

**Institutions and Regulations** of a school choice system have to be set in such a way that competitive pressure is generated and that schools react to this pressure by increasing their productivity. Some examples for regulations are: Which students are allowed to choose schools? Which schools are open to choice? How are places assigned if the demand for a school exceeds its capacity? Does financing follow students?

**Information** has to provide a sound footing for decision-makers (parents, schools, school administration) so that they are able to identify the choices that best fit their preferences.
3. Friedman’s mechanism vs. the real world

In the last chapter I have identified the assumptions that are implicit to the mechanism that was presented by Milton Friedman. Only if these implicit assumptions are met can the mechanism work and can school choice really increase school productivity.

In this chapter I will use evidence from various studies to check how often and how completely these assumptions are met in existing school choice systems. For that purpose, I will break down the crucial assumptions into hypotheses that can be checked empirically.

3.1. Parental school choice behaviour

For Friedman’s mechanism to work, sufficiently many parents of the right children need to identify and choose high-productivity schools.

3.1.1. Hypotheses

Most parents are unfamiliar with the term productivity, at least when referring to schools. And readily available productivity measures for schools are rare, as will be seen in Chapter 3.5. Therefore, productivity is usually replaced in interviews and on school report cards by the expression and concept of “academic quality”. The common measures for academic quality are often -but not always and often only weakly- correlated with school productivity. But these are the measures that parents usually have at their disposal when choosing schools. Therefore, I derive the following hypotheses: For parents to choose highly productive schools, they first need to prefer schools with a high academic quality (1). Then they have to be able to identify schools with a high academic quality (2). Third, the available measures of academic quality need to be positively correlated with school productivity (3).

Finally, if choice behaviour is to change school performance, sufficiently many parents of the right children (4) need to choose high quality schools (5). To clarify points 4 and 5: Most schools prefer to receive bright and disciplined students but are not eager to receive mediocre and disruptive students. If having a high productivity would mean attracting mainly mediocre students, most schools would be less eager to try and raise productivity. And if only a few students actually changed schools, the pressure on schools that results from a “right to choose” would not be high.
It depends strongly on institutions, available information and teacher behaviour, how many and which students would have to actively choose schools to make an impact on school behaviour. I will elaborate on these points in the respective sections below.

3.1.2. Empirical evidence

3.1.2.(1) Parental preferences for schools
When asked in interviews on which characteristics they had based their choice of schools, most parents said “on academic quality” (Goldring/Hausmann 1999, Jacob/Lefgren 2007), “on high scores” (Schneider et al. 1999b) and “mainly on academic quality” (Teske/Schneider 2001). Surveying previous studies, Fossey found in 1994 that “convenience” and “academics” were most often reported in all studies by parents as the reason for their decision (Fossey 1994). According to stated preferences, academic quality is therefore one of the most important factors in choosing a school for parents.

But parents might misstate their motives for choosing schools (Fossey 1994). One characteristic about which parents were proven to be lying is the racial composition of the student body. Race is hardly reported by parents as influencing the choice of schools. But when looking at surfing behaviour on a homepage with school characteristics, a study found that most parents first of all looked at the racial composition of the school (Schneider/Buckley 2002). Another study found that proximity and the racial composition of schools had strongly influenced parental school choices, although the same parents had named neither of these two factors in interviews as having an influence on their decisions (Jacob/Lefgren 2005). If parents misstate the impact of race on their school choice, they might also misstate the importance of academic quality. Therefore relying on what parents say about the importance of academic quality for their choices does not necessarily reveal their true preferences. It is therefore necessary to analyze realized choice behaviour based on hard data to confirm the stated strong parental preference for high academic quality.

Analysing choices made by parents, Fossey found that 25 out of 29 districts that had had a net inflow of students had test scores that were higher than average (Fossey 1994). After the introduction of school choice in the Stockholm metropolitan area, parents mainly switched their children to schools with higher average test results (Söderström 2006). Analyzing the lotteries for admittance to schools in Chicago, studies found that for both elementary schools and high schools, lottery winners ended up in schools with higher previous achievement levels (Cullen et al. 2003, Cullen/Jacob 2007). Studies analyzing house prices found that an
increase in the academic quality measures of schools was followed by an increase in house prices in the attendance area of these schools, indicating the increased demand of parents moving into the area in order to get their children into the local school (for example Jacob/Lefgren 2005, Hoxby 2003a). Analyzing attended and assigned high schools of students in Chicago, Douglas Lauen found that average achievement trumps many of the other variables “… in the Darwinian struggle for explanatory power.” (Lauen 2007a p.16). Thus, based on what parents actually do, I also find that academic quality has a strong impact on school choice.

This impact is however not independent of family background. In the Charlotte-Mecklenburg school district, the school quality was more important for parents with a higher socio-economic status. A study identified two large groups of parents: One group, with a lower average socio-economic status, sent their children to the closest school. The second group cared strongly about academic quality as they perceived it. They searched for high quality schools, reacted strongly to differences in quality and were prepared to accept a long commute in order to reach a school with a high academic quality (Hastings et al. 2006). And according to interview-statements, academic quality was more important for parents who were professionals or from the middle class than for working class parents (Woods 1996). These differences in choice behaviour might be due to differing preferences. Low-income families might for example expect lower returns of education for their children (Hastings et al. 2007). But when parents at randomly selected schools, that were serving low- to middle-income households in Charlotte-Mecklenburg County, were given simplified information about the quality of individual schools, academic performance had a stronger impact on school choice decisions. Receiving this simplified information, that translated test scores into a percentage value and gave the probability for applicants to be admitted at the school, had the same effect as a doubling of the implicit preference for academic performance (Hastings et al. 2007). This evidence implies that the lower observed preference of low-income households for academic quality might in part be due to the fact that these households are less informed about the quality of individual schools.

Summing up, the academic quality of schools is one of the most important characteristics in choosing schools for most parents. It is important to note however, that the presented evidence only identifies the preferences of parents, not the underlying motives. Parents might prefer schools with a high academic quality based on two different motivations. They might want their children to learn as much as possible in school and expect the highest educational quality to be found in schools with a high academic quality. Or parents might want to send
their children to schools with a high prestige for several reasons. They might expect benefits for their children if these have a degree from a prestigious school or because they made friends who are likely to provide good networking opportunities. Parents might also want to share in the prestige of their children who attend a popular school. Or parents might just want to keep their children away from perceived danger an influence from the “wrong” types of peers for various reasons.

This difference in motives does not affect observed behaviour ceteris paribus. But the motives could matter, if the information that is available to parents changes. If a new quality measure would better identify educational quality, parents who want their children to learn as much as possible would immediately choose those schools that have the highest educational quality according to the new measure. Parents who mainly care about prestige would only change their choice behaviour if the ranking of schools by prestige would change as a consequence of the new quality measure. And parents who are mainly concerned about “wrong” influences might only change their choice behaviour once the socio-economic composition of schools has changed as a result of altered choices of other parents.

3.1.2.(2) Information of parents about school quality and other school characteristics

Parents obtain their information about the quality of individual schools mainly from two sources. The first source is officially published information on previous cohorts of students at individual schools. Examples are the league-tables in UK newspapers or school report cards that are frequently used in the USA (Burgess et al. 2004). The second source is word of mouth from friends and relatives (Lauen 2007a, Teske/Schneider 2001, Schneider et al. 1998a). This word of mouth information usually comes mostly from within the same social group concerning socio-economic status and race (Schneider et al. 1998a). Most parents actively seek information about schools before choosing one (Goldring/Hausman 1999) but parents from a low socio-economic background are less active in that regard (Schneider et al. 1998a).

Consequently, the knowledge of parents about potential schools varies across social background. Minority- and low-income parents are less well informed about existing choices and have smaller and less informed networks to tap for unofficial information (Teske/Schneider 2001). The knowledge of parents also varies across characteristics of schools. One study found that, although parents are rather badly informed about school characteristics in general, they do manage to identify schools that are better than average according to those characteristics that they value highly (Schneider et al. 1998a). The fact that
parents consistently pick schools with high academic quality scores indicates that they are able to spot academic quality. By taking socio-demographic characteristics of student populations into account, Lauen manages to show for Chicago that it is indeed the test score achievement of schools that drives parental choices (Lauen 2007a).²

3.1.2.(3) Correlation between common school quality measures and educational school productivity

The most common measures for academic quality are measures that are based on centralized tests. Individual test scores are usually aggregated into school-level measures by computing averages or by determining the share of students that managed to cross certain thresholds. Examples for such quality measures are the percentage of students that meets standards in the Chicago Public School System or on report cards of several US states and the League Tables for English schools.

Wilson and Piebalga (2008) found that the output measures that have the highest impact on parental choice in the UK are only weakly correlated with educational school productivity.³ Looking at realized school choice in Chicago, Lauen (2007a) found that educational school productivity and quality measures based on pure test results were even negatively correlated among the popular schools. At this point I only want to point out that common measures for academic quality represent educational school productivity rather poorly. Chapter 3.5 will elaborate on different types of information about academic quality and their implications.

3.1.2.(4) Choosers and non-choosers

The evidence on what types of children actively choose schools is quite consistent. Goldring and Hausmann (1999) found that white students from high-income families with a higher socio-economic status were the most likely to switch. Analyzing realized high school choices in Chicago, Douglas Lauen found that Hispanics, male students and students from poor families were less likely to apply while students with high test scores and those who had chosen their elementary school were more likely to apply to high schools (Lauen 2007a). Another study found that choosing parents were better educated, better informed about the

² High test scores are highly correlated with an education-friendly socio-demographic background of the student body. Thus it is possible that parents actually desire peers for other characteristics (like middle class background) than for the high test scores that are associated with those characteristics. In that case, the seeming preference for high test scores is only due to an omitted variable problem. Therefore Lauen includes variables describing the socio-demographic composition of schools. Test scores remain among the most powerful determinants of parental school choice in this specification.

³ Wilson created productivity measures for schools by comparing actual student-level test results to predictions based on prior achievement. The correlation between this productivity measure and the measures based on pure test results were in the range of 0.33-0.46.
quality of schools and more likely to be white (Schneider et al. 1998a). When school choice was introduced in some districts of Tel Aviv, students who opted out of their assigned schools had a higher SES background and were more likely to be of the ethnic majority (Lavy 2006). Looking at school entrance lotteries in Chicago, two studies found that applicants were disproportional white and Asian, lived in areas with less poverty and had performed academically better before applying (Cullen et al. 2000, Cullen/Jacob 2007).

According to the existing literature, students who actively choose schools are thus more likely to come from a background with a higher socioeconomic status and higher income. They were academically more successful prior to choice, their parents are better informed about school quality and they are more likely to be from the ethnic majority. Such students are more likely to be academically successful and less likely to cause trouble. They are therefore welcome at most schools.

### 3.1.2.(5) Share of choosers

The share of students that actively chooses schools is often restricted by the regulations of the school choice process and by available capacities. If the share of students that is allowed to change each year is constricted to five percent for example, then at most five percent will attend another school than the assigned one. If the application procedure takes much effort and chances of success are slim, many parents will not even apply. And if at popular schools there are not many open slots that are not reserved for local residents, not many students will be able to switch. Therefore, in order to get an estimate for the share of students that wants to choose schools, it is best to look at school choice systems in which choice is not much constrained and free capacities exist.

After school choice was introduced in Tel Aviv, 38% of students opted out of the school assigned to them within two years of entering a secondary school. Until the end of their school career this figure had risen to 62% (Lavy 2006). When school choice was introduced in a US school district, students were guaranteed a slot at their default school but could list up to three other schools as ordered preferences. Only five percent did not submit a form and between 29% and 63% of parents applied for three alternative schools (Hastings/Kane/Staiger 2005). In the Chicago Public School system, more than half of the students attend a school other than the one in whose attendance area they live (Cullen/Jacob

---

4 The percentage of choosers is represented for subgroups. Of the White students not receiving lunch subsidies, 29% listed three choices. The figure for Black students not receiving subsidies is 55%. For students receiving lunch subsidies, the numbers are 46% for White students and 63% for Black students.
2007, Cullen et al. 2000). In these three school systems, the share of switchers is high enough to put unpopular schools under considerable pressure.

3.1.3. Summary on school choice behaviour of parents
For choice to improve educational productivity, parents have to prefer schools with high academic quality (1). They have to be able to identify the schools with high academic quality (2) and the available measures of academic quality have to be positively correlated with educational school productivity (3). Also, students of types that schools want to attract (4) have to choose schools with a high educational productivity in sufficiently high numbers (5). According to the existing literature presented in this chapter, parents do prefer schools with a high academic quality (1) and are able to identify these schools based on quality measures for previous cohorts (2). It is however not clear, whether the motivation behind the observed preference for a high academic quality is based on prestige, or on the expected educational quality of schools. The actively choosing students are of a type that most schools are happy to receive (4) and if choice is not constrained, these choosing students are numerous enough to create pressure (5). But the academic quality measures on which parents base their choice reflect educational productivity rather poorly (3).

3.2. Behaviour of teachers and schools
For Friedman’s mechanism to work, schools need to react to competitive pressure by raising educational productivity.

3.2.1. Hypotheses
The mechanism assumes schools and teachers to increase educational productivity when they are subject to competitive pressure. For this increase to realize, teachers first need to know how to increase educational productivity, and they need to be able to actually implement whatever is necessary to become more productive (1). Also, schools need to feel enough pressure to change their behaviour. Loosing students initially means only smaller class sizes, which is probably not unpleasant for teachers. Therefore, there need to be some unpleasant consequences of loosing students for the school or for the teachers (2). These consequences could be relocations and lay-offs of teachers at unpopular schools if financing and/or teaching positions follow switching students. Schools that have lost many students could even be
closed. Other unpleasant consequences of loosing students could be a bad reputation and pressure from parents and school authorities.

If competitive pressure induces teachers and schools to change their behaviour, they need to change it in such a way that educational school productivity is increased (3). Another possible reaction would be to try and make a school only more productive in appearance. Such a change in apparent educational productivity could be reached by trying to game the system on which quality measures are based. Or schools and teacher could cheat.

Some critics of school choice fear, that incentives from competitive pressure could crowd out intrinsic motivation (Wilson/Croxson/Atkinson 2004). These critics think that many teachers make an effort for altruistic reasons or because they like to work with children. Introducing high-powered incentives could shift the focus of such teachers. Instead of trying to do a good job, they could focus their energies on improving the features that are measured and rewarded. And an incentive structure that rewards performance on clearly defined features could attract personalities to teaching whose intrinsic motivation is lower. For school choice to improve educational productivity, the intended incentives must therefore not be outweighed by the potential negative incentive effects. Closely related to the incentives issue is the question as to what drives teacher’s actions. How much is the average teacher motivated by altruism and how much is he driven by advancing his own well-being (4).

3.2.2. Empirical evidence

3.2.2.(1) Potential impact of schools and teachers on educational school productivity

The scope of measures for schools and teachers to improve educational productivity depends on the time horizon. In the short run, teachers can work harder, for example by taking more time to prepare lessons or by offering extra classes. Schools can fire and replace the most unproductive members of staff and they can end programs and activities that are not productive (Hoxby 2003b). Schools could for example replace extra-curricular activities with coaching sessions and homework tutoring. Also, schools could focus more on teaching content than on making school “fun” for teachers and students. In the longer run schools could attract better staff by rewarding ability and effort. They could also recognize and abandon pedagogical techniques and curricula that are not productive (Hoxby 2003b). Another way in which choice could result in higher educational productivity in the long run is that productive schools survive and grow while unproductive schools shrink or are even closed down.
Since the scope of measures varies with the time horizon, it makes sense to separate the evidence on educational productivity changes in short and long-run effects of free school choice. I will start with the short run:

In 1999, the UK government introduced a teacher-based performance-related pay policy. This policy relied mainly on measured student outcomes. By using prior achievement data and employing a difference-in-difference approach, the authors found that the scheme did improve test scores and value-added of those students whose teachers were eligible. This improvement was measured one year into the program and averaged half a grade per student, which is equivalent to 73% of a standard deviation in individual student attainment (Atkinson et al. 2004). In 1994 the city of Tel Aviv discontinued the existing busing system in which students were assigned to schools depending on the location of their home. The old system was replaced by a new one in which each student could pick among five high schools. Victor Lavy compared students who were living in a band of 200 to 250 yards from the border of the district to their peers living in a mirroring band on the other side of the district-border. The students in the neighbouring district were not allowed to choose schools but were otherwise identical to the treated group according to all available data and thus were a good control group. By using a difference-in-difference approach Lavy found a sharply reduced dropout rate, a higher matriculation rate and a significant improvement in matriculation-exam outcomes. These effects were even stronger for disadvantaged students. Using an instrumental-variable approach to control for self-selection, estimates showed no difference in the effect between those students who had switched schools and those who had had the right to switch schools but had remained in their assigned school. This indicates that the improvements following the introduction of school choice reflect increases in the educational productivity of schools and are not only due to better matching or to the fact that more students were enrolled in the more productive schools (Lavy 2006). In another study, Lavy analyzed an experiment with performance-related incentive pay in Israel. Teachers of classes in grades 10-12 were eligible and could enter several times, once for each eligible class that they taught. Performance of teachers was measured on a value-added basis.\(^5\) Rewards were sizable and were based on a ranking of teachers in the same group of subjects. The results showed a sizable and significant effect on students in classes whose teachers had entered the

\(^5\) Performance measures were based on the difference between the actual test outcome and a predicted outcome. This prediction was based on a regression that controlled for the students’ background characteristics such as socioeconomic status, the previous school and previous performance.
contest. The improvements appear to come from changes in teaching methods, extra teaching lessons and an increased responsiveness to the needs of students (Lavy 2004).

It is not practicable to identify the medium to long term effects of school choice by analyzing individual school systems. Over time, many things that affect test results change, and after a few years it becomes hard to isolate the various effects. Researchers have instead analyzed cross-section data to determine how educational productivity of public schools varies with competition. Competition to public schools arises from choice options provided by private schools, charter schools or the fact that small school districts make it easier to choose a school by moving the residence of the family.

In the 1990s, Sweden made public funds available to privately run schools that put them financially almost on equal terms with public schools. A study analyzed the resulting variation in the share of private schooling with a cross section, a panel data model and an instrumental variable approach. The authors found that higher competition from private schools lead to a significant and sizeable improvement both in test results and grades in the public schools (Sandström/Bergström 2002). Analyzing competition deriving from vouchers in Milwaukee and from charter schools in Arizona and Michigan, Caroline Hoxby found that higher competition from these alternative schools increased the educational productivity of the affected public schools significantly and sizeably. If all schools in the US faced high levels of competition from private, charter or other public schools, the author estimated that educational school productivity might rise by as much as 28% (Hoxby 2002a).

Whether teachers know how to increase educational productivity is not easy to establish directly. But the fact that teachers react to high-powered incentives (i.e. both rewards and punishments) by swiftly and significantly increasing educational productivity, indicates that the respective knowledge is available.

It is worth mentioning that schools in Tel Aviv and Sweden reacted to the introduction of school choice by increasing educational productivity as soon as the threat of loosing students arose. They did not wait first to see how many students they would loose and how dire the consequences would be. Burgess et al. (2005) observed a similar pattern in England. These observations indicate that schools and teachers try to anticipate the effects of free school choice and adapt their behaviour based on these predictions.
3.2.2.(2) Consequences of loosing students for schools and teachers

The consequences of loosing students vary greatly. Sometimes the loss of students results only in having more funds per remaining student and smaller classes. On the other extreme it happens that schools are closed and reconstituted with a change of management while all teachers are laid off. Although the teachers could be hired again by the new management, they often have to compete on equal footing with outside applicants for their old jobs. How dire the consequences of loosing students exactly are, and whether they are dire at all, depends mainly on the regulations of the school choice system. Therefore, I will analyze the institutional arrangements more closely in the next subchapter (3.3). Here, I will only explore briefly what schools care about if they do not have to fear much from loosing students.

Even if schools do not face any competition, the shame of being labelled as being low-performing could make schools work harder (Ladd 2002). According to evidence uncovered by interviews, principals in the UK care most about how parents view their school. Therefore, they care most about those school characteristics that are important in the view of parents, even if the principals themselves think that these characteristics are otherwise irrelevant. More specifically, principals care most about how their school is viewed relative to neighbouring schools (Wilson/Croxson/Atkinson 2004).

3.2.2.(3) Reactions of schools and teachers to competitive pressure

If schools feel sufficient pressure and incentives are set correctly, they indeed react by trying to improve educational productivity. I have presented some examples at the beginning of this section: When faced with a reward scheme, teachers in Israel responded with changes in teaching methods, extra teaching sessions and by responding more to students needs (Lavy 2004). When school choice was increased in Tel Aviv, this created uncertainty about future enrolment which in turn lead to educational renewal and quality improvements (Lavy 2006). There is also evidence that schools react to competition by rewarding teacher abilities and performance more. Hoxby found that schools under competitive pressure paid higher premiums to teachers for high effort and above average abilities. Math and science skills were also rewarded more than in schools that were facing less competitive pressure (Hoxby 2002b). In reaction to increased competition from charter schools, some schools and districts in the US have tried to improve outcomes by changing curriculum and principals and by empowering teachers (Ladd 2002).

---

6 This is an indirect citation from a paper by P. Heiman and R. Shapira that is only available in Hebrew. The title is: “Parental Choice in Autonomous schools as a Strategy of Restructuring Schooling Systems”
But there are other ways to avoid negative consequences than to really try and improve overall test results by increasing educational productivity.

Schools and teachers have tried to influence who takes part in tests. One way is to reclassify students into categories in which lower goals have to be met. According to anecdotal evidence, schools in the UK encouraged weak students to leave general programs and to change into vocational ones whose results are less important for the reputation of the school (Wilson 2003). Tests, whose outcomes are the basis for assigning accountability measures, often permit special tests for students with disabilities. The “No Child Left Behind” (NCLB) policy of the Bush-administration for example allows additional time for students with disabilities (Figlio/Getzler 2002). Therefore, it might be tempting to assign weak students to special education classes in order to let them take special tests or to prevent them from taking part in the test. In Chicago, the introduction of new accountability measures initially lead to a sudden increase in the share of students with special education status (Jacob 2005). And after the introduction of accountability measures in Florida, schools reclassified low income and low performing students into special education at significantly higher rates than before. These reclassifications were concentrated in schools that were at risk to fail the state standards – and to experience unpleasant consequences (Figlio/Getzler 2002). Another approach that schools took to avoid weak students taking hard tests is to let these students repeat classes (Jacob 2005). These students then either take the test one year later, which provides time for these students to learn more, or they take the test of a lower grade. Some of those students might also drop out of school before they have to retake the test. Another approach is to use disciplinary measures to exclude weak students from tests. In Florida for example, low-performing students are always more likely to be suspended, and are suspended for a longer time than other students. But this difference in the severity of punishment substantially increased during the testing period in which schools were assessed for the NCLB policy. This phenomenon was restricted to students in testing grades, implying that schools used temporary suspensions to exclude weak students from taking the tests. Another possibility to prevent students from taking tests is to have weak students leave the school permanently. In the three years after the introduction of the League Tables\(^7\) in the UK, the number of permanent exclusions of students from individual schools roughly tripled (Wilson 2003). And

---

\(^7\) The League Tables are rankings of schools that are based on the outcomes of central exams and published in newspapers. Parental school choice heavily relies on these rankings.
there is anecdotal evidence that schools in Virginia tried less hard to prevent weak students form dropping out after the introduction of an accountability system (Figlio/Getzler 2002).

Schools have also tried to prevent students who they expected to be weak from enrolling in the school in the first place. When schools were allowed to select their students, they tried to cream-skim or to crop off certain types of students.

Schools that can pick their students have been shown on several occasions to have student compositions that are similar to those of the local public schools that have no influence over their enrolment. But on closer inspection, this picture changes. One study about charter schools on Washington, D.C. found that those charter schools that were market-oriented\textsuperscript{8} enrolled significantly smaller percentages of special education students, students with limited or no English proficiency and students coming from a low-income background. On the other hand, non market-serving charter schools, often sponsored by charitable organizations, enrolled more students with special needs, or who were at risk, than the average public schools. Taking averages of all public schools had masked this difference (Laciero-Paquet et al. 2002). Another study found similar results for autonomous schools in England and charter schools throughout the US. Those schools that handled their own admissions and were profit-oriented tried to cream students who were not costly to educate and who were expected to score well on tests. At the same time there was a second group of schools that handled their own admissions and were founded with the expressed intention of catering to students at risk or with special needs. These schools masked the cropping behaviour of the first group in comparisons between schools handling their own admissions and public schools (West et al. 2006). In all these cases, schools were in theory required not to discriminate in the admission process and were supposed to achieve a student composition that resembled that of the neighbourhoods surrounding their location.

A study on charter schools in California sheds some light on the methods employed to affect the student composition (Wells et al. 1998). These charter schools used for example a diverse set of means to make students and parents aware of their existence and their programs. These means included brochures, mailers, presentations at public forums, advertisements in local newspapers and word of mouth. Frequently, these means were targeted at desired student groups, for example based on geographical location of residence or language proficiency.

\textsuperscript{8} The authors classified schools as market-oriented based on seven characteristics that mostly checked whether for-profit organizations were involved in founding or running the school and whether the school had multiple campuses or was planning to expand.
A large share of the Californian Charter Schools also used interviews to assess the match between the school policy and prospective students. Based on information gathered during these interviews, students were steered to apply or not to apply. Several factors were named as relevant for whether a student was encouraged to apply by the school administrations. Among these were behaviour, effort, previous academic achievements and the parental ability to get involved. Apart from active screening by schools, such pre-admission interviews require an effort and a willingness to undergo this procedure and may thus be used to discourage applications of students who the school desires to avoid (Wilson 2003). Extensive and complicated application forms and application procedures in general may also be used to screen the potential pool of applicants for those who are willing to make an effort and are capable of correctly handling the process (Propper/Wilson 2003).

Transportation also plays a role in screening parents. The charter schools in California rarely provided transportation to school. This places students at a disadvantage who are living farther away or whose parents are either not willing or not capable of ferrying their children to school every day (Wells et al. 1998).

Some schools and teachers don’t stop at trying to influence who takes part in tests but also try to influence the test outcome through various other measures. If quality measures contain information about the percentage of students that has passed certain thresholds, schools have focused their attention on those students on the borderline of meeting these thresholds, often at the expense of other students (Propper/Wilson 2003). When tests are repeated in a similar form or if schools and teachers know what to expect, they can try to “teach to the test”. This means that teachers focus heavily on the expected content of the upcoming test, neglecting other parts of the curriculum. Brian Jacob found that the introduction of accountability measures connected to the “No Child Left Behind” policy in Chicago lead to sharp increases on the high-stakes test on which the measures were based. But it did not lead to increases in a broader, state-administered low-stakes exam. This indicates that teachers had strongly focused on preparing their students for the high-stakes test which lead to significant increases in this test only (Jacob 2005). If teachers knew the actual questions before the test took place, they sometimes actually taught the test, meaning that they let the students do the test some time before the real one. When teachers administered the test to students of their own school, some gave the correct answers during the test. In one testing regime, teachers were not in the classroom during the test but were given the task of “cleaning” the answer keys. They had to clean stray pencil marks, remove dirt and darken
item responses that were only faintly marked by students. This task was necessary so that the
answer forms could be processed by machines. But by letting teachers do the “cleaning” of
their own students, they were given an opportunity to cheat. By using clever econometrical
techniques, researchers found that at least 4-5% of teachers had “corrected” the answers of
their students in order to achieve higher scores (Jacob/Levitt 2003). Some schools even tried
to improve test scores by feeding the students more calories. By analyzing school menus, a
study found that schools in Virginia that were threatened with sanctions from accountability
programmes tried to improve test results by increasing the calorie content of school meals on
testing days (Figlio/Winicki 2005).

3.2.2.(4) Altruistic vs. selfish teacher motivation:

If teachers were mainly interested in the academic achievement of their students, it might be a
good idea to give them a free reign in schools and there would be scarce need for further incentives. There is evidence that teachers are indeed to some extent intrinsically motivated (Wilson/Croxson/Atkinson 2004). This intrinsic motivation might entail that teachers enjoy teaching successfully, that they care for the future well-being of their students or simply that they like doing their job and thus do it enthusiastically. Outcome-based incentive systems usually work either directly through sanctions and rewards or indirectly through choice and competitive pressure. Forcing such an incentive system on intrinsically motivated teachers might have undesired consequences. Teachers might respond by focusing only on those tasks for which output is measured. And if teachers think that the performance measurement that underlies sanctions and rewards, or that is used to provide quality information is unfair, they might get the impression that they do not have a chance to benefit from working hard. In this case, performance incentives might even undermine motivation and reduce educational productivity.

It is practically impossible to disentangle these effects from the intended incentive effects of sanctions and rewards or from effects working through pressure from informed parental choice9. But for this study only the net effects are of interest: Does the increase of pressure

---

9 A negative incentive effect would be opposed to the intended effect of performance incentives. Only the aggregate can be observed. It might be possible to observe isolated negative incentive effects under extreme circumstances. A school that has a weak student intake and is on probation under rules like in the NCLB policy would be a possibility. Such a school would have to improve test results greatly to avoid punishment. But such improvements are sometimes impossible given the student intake. If teachers knew that their effort could not prevent punishment, there would be no positive incentive effects of the increased pressure and unwanted incentive effects could be watched in isolation. But such estimates of negative incentive effects would apply only to the extreme situations in which they can be measured.
induce the average teacher in a typical school system to increase his effort or not? If pressure leads to such an increase in effort, then the marginal effect of pressure is an increase in educational productivity.

Caroline Hoxby analysed the influence of teacher unions by exploiting external legal settings that helped or delayed the rise of teacher union organisation and their accumulation of power. She found that if teacher unions had more influence, teacher pay increased and test scores deteriorated (Hoxby 1996). This suggests that if teachers can influence their working conditions, they try to get more money and try to expend less effort. This pattern indicates that the average teacher is on the margin more interested in advancing his own well-being than that of his students.

Increased performance incentives in Israel, Sweden and the UK that were described above resulted in an increased educational productivity. This suggests that at the margin the intended incentives of increased competition and/or performance-based accountability measures outweigh any negative incentives.

3.2.3. Summary concerning the role of schools and teachers in school choice

For school choice to increase educational productivity, schools and teachers first have to know how to increase educational productivity. And they have to be able to achieve such increases (1). There have to be negative consequences of loosing students and/or positive consequences of attracting students for schools or teachers (2). If schools respond to pressure that results from school choice and its consequences, this reaction has to lead to an increase in educational productivity (3). Also, for the intended incentives of school choice to lead to increased educational productivity, these incentives must not be outweighed by accompanying unintended incentives (4). And the preferences of the average teacher have to be shaped in a way that leads to more effort in the face of increased pressure (5).

As can be seen from the presented literature, teachers know how to achieve a sizable increase in educational productivity and are able to reach it, if given the right incentives (1). The consequences of loosing and attracting students depend strongly on regulations that I will analyze in the next subchapter. In the absence of direct consequences, teachers and principals care most about how parents see the quality of their school relative to other close-by schools (3). The effects of unintended incentives are hard to observe, but increased competitive pressure leads to swift and sizable increases in educational productivity. This response to competitive pressure and the observations of Caroline Hoxby suggest, that the average teacher cares at the margin more about increasing his own well-being than for increasing the quality
of the school for the sake of the students (4). Summing up, the average teacher has preferences that lead him to react to increased pressure by trying to increase educational productivity (5).

3.3. The effects of institutions and regulations on school choice systems

The impact of school choice on educational productivity heavily depends on institutions and regulations. These need to be set in such a way, that school choice can generate competitive pressure. And institutions and regulations have to ensure that actors respond to this pressure in ways that increase educational productivity.

3.3.1. Hypotheses

Many regulations can constrict choice in a school system if set in the wrong way. If, for example, only a small fraction of the students is allowed to choose schools each year, then only a few will leave their attendance area school. A scarcity of free slots in the school system, or the earmarking of free places to certain groups of students, can result in a situation where most students will not really be able to choose schools. When there are only a few schools within commuting distance of most households, as it is common in sparsely populated areas, active school choice will also be rare. And a restriction of eligible schools to only a few for each student can also constrict choice.

If there is little potential for school choosers, either because students are not allowed to leave or because there is nowhere they can go, then schools do not have to fear to loose many students. In that case, competitive pressure will be low. Thus, for school choice to create competitive pressure, there has to be sufficient potential for school changes (1).

Regulations need also to be set in a way that encourages applications for school choice (2). If everybody has to fill out school choice forms, there will be more competition than if everybody is assigned to a default school and can apply to schools of choice, which then usually requires initiative and additional effort. And if parents only have to insert the name of desired schools into forms, there will be more competition than if the application to schools is complicated and time-consuming and has to be repeated for each school to which the student wants to apply.

For free school choice to lead to competitive pressure, regulations have to be set so that it is beneficial for schools and teachers to attract students (3). One way to achieve this is if financing and/or teaching positions follow the choosing students. Then, teachers and
principals will fear for their jobs and will likely respond more strongly to competition. If financing follows students only partially, loosing students means having more funds per remaining student. And if financing does not depend on the type of the student, it might be good for schools if costly student types leave. Therefore it is important how much financing follows each type of choosing student relative to the costs of educating her. In some school systems there are additional sanctions and benefits attached to changes in student numbers and performance measures. If schools that loose too many students are closed down, incentive effects from school choice should be higher than if contracting schools get special assistance.

Whether schools want to attract students also depends on the types of students that are allowed to choose (4). If only disadvantaged students, underperforming ones or those from failing schools are allowed to choose, schools will likely invest less effort to attract them than if all students are allowed to choose. Which types of students the schools want to attract depends on financing schemes and school evaluation procedures. If receiving a costly student also means getting more funds than the average student cost, schools will be more willing to attract such a student. And if schools are evaluated based on student progress, they should be more willing to attract underperforming students than if evaluation is based only on average test results.

Another group of regulations that influence whether schools attempt to attract students determine which student ends up in which school (5). One way to regulate student assignment is to allow schools to pick among the applicants. They could be allowed to pick based on previous grades, based on other defined and verifiable rules or they could be allowed to pick freely at their own discretion. Schools could be allowed to reject unwanted students generally or they could only be allowed to pick students if the number of applicants exceeds the number of free slots. Alternatively, student assignment could be handled centrally. One common way to assign students is to hold lotteries among applicants for over-subscribed schools. These lotteries could be held separately for each school or they could try to assign each student to one of all his (ranked) choices. The assignment mechanism could aim to result in a desired mix of students’ backgrounds or races. And there could be preferential access for local residents or for students with siblings who already attend the school. The more discretion schools have in choosing their students, the better it is for them to attract applications. But more discretionary power of schools could also lead to cream-skimming. If schools were allowed to pick only the best students, then free school choice could lead to a segregated school system in which the best students accumulate in a few schools.
Which types of students the schools want to attract also depends on quality measures that are used to evaluate schools (6). If quality measures based on average test results are used, schools will try not to attract students that they expect to be underperforming. If quality measures are based on value added or on student progress instead, these students might be more attractive. It is also important whether and which quality measures are publicly available so that parents can base their choice on them.

There are also cross-effects of regulations and institutions (7). Some of the regulations that were described above affect the outcomes of other regulations. I have already presented some of these interactions and it is not possible to analyse them exhaustively in this dissertation. But I will briefly present two sets of regulations that are not uncommon in school choice systems and that are unfavourable to the mechanism that was presented by Friedman:

If average test results are published, or used to assign sanctions and rewards, schools will not want to attract students who they expect to be underperforming. If additionally schools are not allowed to pick or reject students, they will try to influence which students apply to them. If only underperforming students or those from failing schools are allowed to choose, schools will expect most applicants to be academically worse than average. To receive additional students would under these regulations mean a decline of average test scores. Thus, schools should not be expected to make an effort to attract students in this case. Consequently, educational productivity increases due to school choice would be surprising in such a setting.

Example for this type of setting are programs that aim to help disadvantaged students to get into better schools, such as income-targeted vouchers or the income-targeted right to attend schools in other districts.

Another combination of regulations that is unfavourable to the mechanism presented by Friedman is the following: School quality measures are based on average test results. There are sanctions and rewards attached to these measures, if only by loosing students and thus funding and/or staff positions. All students can apply freely but schools handle application procedures. In this case, schools do not want to attract underperforming students but are happy to receive good ones. If schools can pick students, they will only accept the best ones. This is likely to create a self-reinforcing circle: “good” schools receive many applications, pick the best students and therefore score even better on tests next year. This can happen even if these schools are not good at teaching at all, but only happened to have a good student intake in the previous period. If on the other hand schools are not allowed to pick, they might try to screen students through complicated application procedures or tailored programmes. Complicated and time-consuming applications help to deter or even exclude parents who do
not care much about school quality or for whom the forms are a challenge. A weak support for special education students or those with a foreign mother tongue might deter these types of students, while a focus on ancient languages or on sports that are popular in the upper classes might attract wanted types of students. If schools can influence, either directly or indirectly, which students they take in, this might lead to heavy sorting. The outcome might be a distribution of students where high performing students accumulate in a few choice schools that are usually situated in prosperous neighbourhoods. Examples for this type of choice system that I have mentioned before are charter schools in Washington, D.C. and autonomous schools in England.

3.3.2. Empirical evidence

The findings in the previous sections were the outcome of preferences that are similar for most parents or teachers and of mechanisms that hold in general. Contrary to the previous sections, there are no general findings to be presented in this one. Each existing school choice system has its own particular set of institutions and regulations. This set is the result of political processes. The actors in these processes are driven by quite different motivations and the outcomes of these processes are influenced by power distributions and decision mechanisms which vary across school constituencies\(^{10}\). The results are therefore diverse. I will thus not be able to give general results. Instead, I will present the diverse specifications of regulations that can be observed in existing school choice systems and give some information about how common these specifications are.

3.3.2.(1) Potential for school choice to create competitive pressure

In some school choice systems, students are allowed to choose schools quite freely. In Tel Aviv, each student can choose among 5 schools, 3 of them inside his own district and 2 in other districts (Lavy 2006). In the Chicago Public School system, students can apply to most schools and are allowed to apply to as many schools as they like (Cullen et al. 2000). In England, students can choose any school within their local education area, provided that it has free places (Burgess et. al. 2004). And when school choice was introduced in the Charlotte-

\(^{10}\) School choice might be designed to maximize productivity gains. The design might also be intended to result in a desired mix of different socio-economic classes, ethnicities or races at each school. School choice might be introduced to help certain types of students specifically or it might be set up to give every student the same opportunities. Teacher unions might be more or less influential and might try to avoid uncertainties that arise from student migrations between schools. Decisions might be taken by elected school boards that include administrators, teachers and parents in diverse proportions like in the US or by regulating bodies staffed exclusively with bureaucrats like in many European countries.
Mecklenburg School District, parents could hand in a list of three ranked choices among all schools in the district (Hastings/Kane/Staiger 2005). But often, school choice is more limited. When vouchers were introduced in Milwaukee for example, they were initially only available for poor students (Hoxby 2003b).

The number of students who are allowed to choose schools each year varies across school choice systems. All students are free to apply to schools and to switch if they find a school that accepts them in Chicago (Cullen et al. 2003), Tel Aviv (Lavy 2006) and England (Burgess et al. 2004). The only constraining factor in these systems is the availability of free places at receiving schools. In other school choice systems, the total number of school choosers is limited. This is most likely the case if choice crosses financing boundaries, like in inter-district school choice or choice from public into charter schools in Michigan (Hoxby 2002a) or via vouchers from public schools into privately run ones. A voucher program in Milwaukee was for example limited initially to one percent of total student enrolment (Hoxby 2003b). And privately financed programmes in New York and Washington D.C. provided vouchers only to 1300 and 460 students respectively (Ladd 2002). If the number of allowed choosers is constricted to such small fractions, there will not be much competitive pressure.

The capacity of free places at schools also varies across school choice systems. It was possible for 38% of the students to change schools within two years after the introduction of choice in Tel Aviv (Lavy 2006). And in Chicago there is sufficient excess capacity, partially created by closing and reconstituting schools, that one in two students goes not to the school in which attainment area she lives (Cullen et al. 2000). But in other school choice systems, a scarcity of free capacities stifles competition. In England, students can apply quite freely to public schools and about half of the students do not attend the nearest school. But places at schools are earmarked to local residents and to those whose siblings already attend the school. These regulations effectively turn some of the most popular schools in England into neighbourhood schools, making it almost impossible for outsiders to get in (Whitty 1997, Burgess et al. 2004). One study identified the lack of free places at many schools as the most important constraint for school choice in England (Burgess et al. 2005). Similar regulations seem to have stifled school choice in New York in 2001 (NYT 2001). If the best schools are effectively turned into closed shops, the attractiveness of their high academic quality to students does not pose a threat to other schools. Even if most of their students wanted to choose the best schools, they would not be able to get in.

In urban areas like London, Chicago or Tel Aviv, there are usually several schools within easy commuting distance. But in sparsely populated areas there is often no real school choice
due to geographical reasons. In London, there are for example 17 schools within a 10 minute commuting distance of the average household, but in rural areas this figure goes down to about one school (Burgess et al. 2004).

Summing up, the potential for school choosers depends heavily on regulations and geography and varies across school choice systems from vast to virtually non-existent.

3.3.2.(2) Procedures for applications and student assignment

When free school choice was introduced in the Charlotte-Mecklenburg School District, parents were guaranteed a place at a nearby school and could hand in a list of three ordered school choices. This application procedure, which took not much time and effort, lead to only 5% of parents who did not submit a form, 95% did (Hastings/Kane/Staiger 2005). In Tel Aviv choice was obligatory. Students did not get a guaranteed place at any school. They had to rank the five schools available to them. Then each student was assigned to one of these schools. In the first years of the program, assignments were based on student priorities and the goal to have a socio-economic balance at each school that matched the balance of the city. After 2003, places were assigned by lottery (Lavy 2006). And in Chicago, applying to a school does not involve much effort and parents are likely to apply to more than one school. Cullen et al. (2003) analyzed applications to high-schools in Chicago. Although they only used data on less than a third of the regular high schools, they found about 19,500 applications by about 14,500 students.

Applications in targeted choice programs can be more complicated. Parents have to prove first that they qualify. If the right to choose schools is limited to the poorest of households, this means for a household to reveal, and prove, that its income is low. Some parents might hesitate to reveal their financial situation in this way. Schools that were not allowed to reject unwanted applicants openly but were in charge of the application procedure, tried to dissuade unwanted students by using tailored application forms (Propper/Wilson 2003). Parents form minorities, with underperforming children and from low SES-backgrounds are less likely to apply even if application procedures are rather simple and easy to file. The necessity to deal with application forms that are made complicated on purpose, and often in a foreign language, could prevent many parents with an uneducated or migration background from applying.

Another factor that affects the likelihood of parents to apply is the probability with which they expect their child to get into the school. As schools reduced class sizes and the number of

---

11 See 3.1.2.(4)
slots open to applicants from outside their district declined in the late 90s, the number of applications dropped in New York (NYT 2001).

Summing up, regulations regarding application procedures vary from settings that force all parents to rank their preferred schools to settings where students are assigned to schools and where applying to other schools is voluntary, complicated and time-intensive, and thus requires considerable initiative and effort.

3.3.2.(3) Consequences of attracting or loosing students

The consequences of attracting and loosing students vary greatly. If many students leave, teachers might be laid off or transferred. This might happen directly through orders from the school regulators. Alternatively, financing follows leaving students and schools have to react by changing wages or the number of staff. In extreme cases of student loss, school closure could be a consequence. A common reason for school closure in the Chicago Public School system was that the facilities were underused after too many students had left (NCBG 2007). Sometimes these schools were reconstituted under a new leadership, often run by private companies. But in any case, teachers lost their positions. They had to apply again to the new leadership or had to be transferred elsewhere in the district (Jacob 2005). And in the first years of the school choice program in Tel Aviv, one of the nine high schools was closed (Lavy 2006). Loosing ones job or having to go to another school while bearing the stigma of coming from a failed school should be a powerful motivator. But often it might not be possible to close down schools, no matter how bad they are. If there is a capacity-shortage and compulsory schooling, it might not be possible to close down schools (Ladd 2002). Regulations sometimes consider capacity-shortages. According to guidelines of the Chicago Public Schools for example, a high school can not be closed if it is not possible to accommodate all its students at other schools that are within walking distance (CPS 2005). And if school regulations do not allow privately run schools, and teachers cannot be laid off, teachers have not that much to fear from competition.

One example of a school choice system in which financing follows students almost completely are charter schools in Michigan. If a charter school attracts another student, it gets financing roughly equivalent to per pupil spending in Michigan public schools. And the school district in which the student lives looses about the same amount of money (Hoxby 2003b).

In other school choice systems, the financing that follows students is less than the cost of educating them. Sometimes no money follows at all. Low transfers of funds are common for
vouchers in the US. Such vouchers provide families with money that they can use to pay for part of the tuition fees at a choice school. The typical voucher in the US amounts to only 14 – 29 percent of per-pupil expenditure in the local public schools. And in the early stages of a voucher program in Milwaukee, the state even ensured that schools that lost students would not lose any financing (Hoxby 2003b). Another type of school choice in which financing does not follow students completely is inter-district school choice in the US. There, a large share of school funding is financed by local property taxes. Under such a financing regime, a student who leaves with financing for a school in another tax area or who leaves the public education system, effectively takes money from local taxpayers and gives it to another district or an entity that is not state-run. Local tax payers, schools and teachers and the school district authorities dislike such an outflow of funds. These groups are usually better organized than potential beneficiaries of free school choice across districts. And together, they have more political influence. Therefore, usually only a small fraction of financing follows inter-district school changers, like in Minnesota (Hoxby 2003b).

Even if financing comparable to average education costs follows students, this might not be enough. Special education students are for example often more expensive to educate than other students. And language assistance for students who are no native speakers also is costly. If only average costs are transferred, receiving a special education student means losing money on a per student basis. Such financing regulations might make schools cut down special education programs in an attempt to discourage students who need them.

Additionally to indirect consequences from losing students, school authorities tie sanctions and rewards directly to performance in many school systems. In some school systems, badly performing schools not only risk losing students, but can get closed down because of bad performance. In Chicago, failing schools are set on probation and then, at least if there are slots for their students at nearby schools, might be closed down if they do not improve (CPS 2005). In Florida, students of schools that got the worst possible grade in the accountability system of the state twice within four years got vouchers which they could use to go to private schools (Ladd 2002). The No Child Left Behind policy forced all US states to create an accountability system. If students do not make a yearly progress that is deemed adequate, schools and districts face consequences. Among these are mandatory public school choice or even complete school restructuring. Also federal funds could be redirected, meaning that schools, districts or complete states could lose federal financing (Figlio/Getzler 2002). Some US states have a system of sanctions and rewards. Bad performance is punished by the loss of money, restructuring or even closing of schools. On the other hand, districts, schools or single
teachers can get rewards for good performance, often based on progress, not on average test results. These rewards can be substantial (Kane/Staiger 2002). Although the assignment of sanctions and rewards often seems at least somewhat arbitrary due to measurement-problems (more on that in the next section), such measures can provide powerful incentives. These incentives often complement, but sometimes counteract the incentives that arise from competition induced by school choice. As a consequence, measuring the effects of school choice in isolation becomes more complicated. Especially if, as it often happens, failing schools get special support in the form of consultation, additional teaching staff or a changed school leadership.

3.3.2.(4) Characteristics of school-choosers and attractiveness of choosers for schools
In many school choice systems, most or all students are allowed to choose schools. Examples described so far include England, Chicago, Sweden and Tel Aviv.

But the right to choose schools is not always universal. Especially voucher programs are often targeted at poor students or at students of failing schools. The voucher program for students in Milwaukee was, for example limited to children of families with an income below 175% of the federal poverty level (Hoxby 2003b). And in Florida, only students of failing schools were given vouchers (Ladd 2002).

If school choice is limited to poor students or those of failing schools, it is likely that schools expect choosers to be academically weak.

If school choice is voluntary, most choosing students are of types that most schools want to attract, as shown in 3.1.2 (4). If school choice is limited to poor students or those of underperforming or closed down schools, the choosers are likely to be of types that most schools do not want to attract. Which types of students are wanted or not wanted by schools, also depends on other regulations. These include: (a) the measures used to gauge the quality of schools, (b) how students are assigned to schools and (c) how much financing follows each type of student. But in most cases it is more attractive for schools to get many applications if everybody is allowed to choose, than if only disadvantaged students are allowed to choose.

3.3.2.(5) Assignment of students to schools
In many school systems in which choice is possible, some schools decide which students they take in. Privately run schools are often allowed to pick their intake. Sometimes this selection happens indirectly, via tuition fees. Only parents who are rather wealthy, and care enough about the education of their children to pay for it, can send their children to most of the private schools. Many privately run schools also select their intake directly. Such selection is
often based on prior achievement and conduct. But selective intake is not limited to private schools. Sometimes public schools are also allowed to choose their intake. All-male or all-female schools obviously base their intake on gender, and specialized schools only admit students who belong to that group of students to whom the school caters. Such groups could be students with certain needs in special education, especially bright students or those who are musically gifted.

If schools can influence their intake, they might be tempted to pick those students who are easy to educate or will help the school look good. If such a selection is combined with quality measures based on average test results, it can create a virtuous circle for individual schools: a good intake leads to good test results which again attract a good intake. No external energy (like making an effort to provide a better education) is necessary to keep such a circle running. There is evidence that suggests that this kind of circles based on cream-skimming exists. In England for example, schools that were judged to be good based on the pure output measures and thus had ample applicants to pick from were most likely to be academically selective or formerly selective, meaning that the average accepted students had above-average scores in previous grades (Whitty 1997). West et al. (2006) found evidence that suggests cream-skimming for charter schools in the United States and autonomous secondary schools in England. The introduction of new value-added measures provided a new perspective in some school choice systems. Schools that ranked highly based on outputs generally had an average ability of their students in the highest quintile and thus owed their good results mostly to a good student intake. When regarding value-added measures, the best schools came from a wide spread of average student intake ability (Wilson 2003). This means, that the best scoring schools according to output measures were not much more likely to be better at actually educating the students. But as parents cared most about the pure output-measures, these schools were over-subscribed. And as local residents are accepted first in England, schools that were situated in areas with a good intake can consistently get good test results and can therefore take in students almost exclusively from the good local pool again, even with school choice.

Even if schools themselves are not allowed to pick students, there are many regulations that can constrict free choice. As mentioned above and in 3.3.2.(1), local residents or applicants whose siblings are already attending the school are more likely to be admitted in many school choice systems. In this case, entrance to popular schools depends heavily on home location. Another set of regulations that affect student admission are rules that aim at a certain mix of
ethnicities, races or social groups. Before the introduction of school choice in Tel Aviv for example, students were assigned to schools based on their home location. And they were carried to their school with buses to make sure that each school had the intended mix of students by income and ethnic background (Lavy 2006).

But there also are school choice systems where school assignment is based on none or at least hardly any student characteristics. In Charlotte-Mecklenburg School District, student assignment was only based on ranked student choices and school capacities (Hastings/Kane/Staiger 2005). In most of the Chicago Public Schools, free slots are assigned to groups based on race and sometimes based on social background. Then there is a lottery held for each student group that distributes the available places among all students who belong to this group (Cullen et al. 2003). And in Tel Aviv excess applications were handled by a lottery from 2003 on (Lavy 2006).

3.3.2.(6) Effect of school evaluation procedures on attractiveness of choosers for schools

Which types of students are attractive for schools depends heavily on the type of quality measures that are used in the school system. The different ways to generate quality measures and to present the resulting information to parents are dealt with in the next section (3.4.).

At this point I only want to summarize some insights from previous sections. Generally, if quality measures have consequences, either directly through rewards and sanctions or indirectly through parental school choice, it is likely that schools and teachers try to score well. This reaction is one of the main reasons put forward by proponents of school choice and is the cornerstone of Friedman’s mechanism. But if the regulations provide schools and teachers with loopholes that make it possible to improve outcomes in ways other than actually providing a better education, it is likely that some schools will be tempted to exploit these loopholes.

If, for example, not all students have to be accounted for, schools or teachers might try to keep weak students out of tests by encouraging them to stay at home at the testing day or to drop out entirely. If students in special education are exempt from tests or are permitted to take special tests, schools or teachers might try to reclassify weak students into special education classes. But if students in special education do not receive a special treatment, schools might try to attract as few of these students as possible. If average test outcomes are used that do not account for student intake, schools might try to prevent weak students from signing up. Schools could simply refuse weak students if they are allowed to do so. Otherwise they could try to dissuade them from applying with complicated application procedures,
application interviews or a school concept that does not cater to needs that are typical for students who score badly in tests. If sanctions depend on the share of students that fails to meet given standards, schools and teachers might focus their effort on those students who are close to meeting the standards. And if consequences depend mainly on a few subject areas, schools and teachers could focus on these areas, neglecting others.

Summing up, it is necessary to design an accountability system carefully. In doing so, an administration should think about ways in which teachers and schools might respond to the given incentives in ways that differ from those which the accountability system is intended to induce. It might be necessary to take additional precautions in order to prevent unintended reactions. A school administration might, for example, want to use special testing procedures for students with disabilities, both to get unbiased estimates of school outcomes and to avoid a situation in which schools try hard to scare away students with disabilities. But if such special testing procedures also help other students to score well -like more time to answer the questions- some schools or teachers might try classify students without disabilities as disabled. A possible precaution against such false classifications could be, to have decisions about a disability status to be made at that administrative level which also measures school quality.

3.3.3. Summary on the effects of institutions and regulations

The specifications of institutions and regulations in school choice systems are quite diverse. And these regulations heavily affect the process and the effects of school choice. Institutions and regulations determine, for example, how much scope there is for school choice, from virtually none to the real possibility for each student to go to a school of her choice (1)\textsuperscript{12}. These specifications also influence which and how many students apply to other schools (2) and whether the loss of students leads to consequences anywhere between only teaching smaller classes and the threat of school closure and lay-offs of the entire staff (3). The specifications of institutions and regulations also influence what types of students apply to schools (4) and to which schools they are eventually assigned (5). The type of school evaluation is also determined by regulations (6). Finally, there are many cross-effects of individual institutions and regulations that are too numerous and complicated to be presented here in full but that can easily break Friedman’s mechanism, as shown in two examples (7).

\textsuperscript{12} These numbers refer to sub-sections 3.3.1.(X) which presented the theory and 3.3.2.(X) which presented findings in the previous literature on the discussed issue.
Insights into effects and cross-effects of institutions and regulations are so far patchy at best. And they are not easy to achieve for reasons that I will discuss in chapter 4.

3.4. Students and effects of changing schools

So far, I have only analyzed schools, teachers and parents. But it is the students who change to other schools on whom this change will have a direct impact. In this section, I will analyze the preferences of students regarding school choice and how they are affected when they change the school that they are attending.

3.4.1. Hypotheses

Parents are legally responsible for making school choice decisions, but it is the students who actually change schools. So far I have looked at the preferences of parents only. Although I cannot identify whether the underlying motivation is based on educational quality or prestige, the resulting parental preferences are clear. For most parents a high academic quality has a high priority when choosing schools.

But when students enter high-school, a good education or a prestigious degree and the resulting effects on higher education and the job market might not be their primary concern. Students might therefore have different preferences than their parents and a high academic quality might be less important to them. Instead, students might want to go to schools with other characteristics, such as good facilities for their favourite sports, a reputation for not being too challenging academically or being attended by most of their friends. Therefore it is important to know, who actually makes the school choice decisions. Is it the parents who choose the schools, is it the students or is this choice made by parents and students together? If students do have an influence in the decision, it is important to know their preferences (1).

It is also important to know what happens to students who change schools. Central to Friedman’s mechanism is, that parents choose the more productive schools. Only if a higher educational productivity attracts more students can it be worthwhile for schools to try to increase it. This incentive is what drives the increase in overall educational productivity that is expected to result from school choice. If students choose the more productive schools, as assumed, they should gain academically. It should therefore be possible to find empirical evidence for such gains (2). Additionally to a change in the quality of academic instruction, changing the school might have other effects. Students who change from an assigned school, which is usually close to their home, to a school of their choice will likely have to commute
farther to school. Commuting costs time and is tiring. It might therefore have an effect on the well-being of the students and on their academic success. And concerns of parents about the costs of commuting could influence their school choice decision (3). Additionally, there might be other effects of changing schools. Students who change to another school will have new peers and teachers, will be in an unknown environment and will probably be separated from their old friends. This might have negative initial effects on satisfaction with the school and on academic outcomes (4). Another possible result of changing schools could be a better match between preferences and needs of students, or of their parents, and the characteristics of the attended school, resulting in higher levels of satisfaction and better academic outcomes (5).

3.4.2. Empirical evidence

3.4.2.(1) School choice by households

I have not found any representative and reliable information about preferences that students have about schools. But there is evidence that the preferences of students do not have a lot of influence on the school choice decision:

When interviewed, 77% of the parents stated that they had taken the school choice decision on their own, another 22% stated that they had taken the decision together with their children (Woods 1996). According to these statements, only 1% of the students could choose their school alone and ¾ of the students were not involved in the school choice decision. The size of these figures should be interpreted with caution, however. As with all interview data, these findings could be distorted deliberately, or by false memories or perceptions. Parents might state that they had included the child in the school choice decision because such cooperation fits a popular education style. But in reality, they might have presented their child only with a selection of schools with high academic quality. The parents might also have presented the features of schools to their child in a biased way. The misrepresentation of the decision process could also be the other way round. Parents might state that they have taken the decision alone, but in doing so they might have cared more about the current preferences of their child than about her future prospects on the labour market or about their own preferences. But even when considering some uncertainty about the truthfulness of statements, the size of these figures indicates that parental preferences dominate in the school choice decision over potentially differing preferences of the students.

Evidence based on realized choices showed in Section 3.1.2.(1), that the most important drivers of school choice decisions were proximity and perceived academic quality. These are
the characteristics of schools that are most important to parents. Summing up, students either have similar preferences to those of their parents or they do not have a strong influence on the decision. In either case, parental preferences can be used to determine school choices of households.

3.4.2.(2) Academic gains as a result of school change

The grading systems or standards of individual schools might differ. Therefore, only the results of standardized tests should be used to analyze how the academic performance of students changes when they change schools. Most researchers are convinced that some schools are better at educating students than others and therefore expect to see clear effects of at least some school changes on the test outcomes of individual students. But the strand of literature that analyzes these effects of school changes on academic outcomes has produced unclear and conflicting evidence. Before presenting any study in detail, I will address some common features of school choice systems. These features make common approaches to investigating academic effects of school changes difficult or even futile.

If school choice is voluntary, we have seen that the students who apply to other schools differ systematically from those who do not try to actively choose a school. Section 3.1.2.(4) showed that choosing students are usually of types that can be expected to score better than average on tests. A simple comparison of school-changers to non-changers would therefore be partly driven by systematic differences between the two groups of students that are due to self-selection into the treatment.

If schools can pick their students, differences in test results might be largely due to another type of selection bias. In this case, oversubscribed schools would likely pick those students who they expect to score well in tests. A simple comparison between school-changers and non-changers would then tell more about the preferences of schools regarding student types than about the effects of choosing schools on students.

Student assignment policies can also distort estimates. If local residents are admitted with priority, popular schools often are in areas with a good local student population and can turn effectively into neighbourhood schools (see 3.3.2.(5) ). Most realized school choices will then be made by rather weak students and into less popular schools.

Another problem would be, if many students change schools for reasons that are not related to the quality of schools. One example can be found in the inner cities in the USA. For example in the CPS school district, that covers the inner city of Chicago, most schools have yearly mobility rates in the double digits, often above 40% (CPS School Data). This is mainly due to
an unstable low-budget renting market and unstable incomes of the parents. According to a
CPS official who is responsible for data management, around 70-80% of the students stay in
the school to which they were admitted at the beginning of the school year. The remaining 20-
30% move around frequently, often changing schools several times a year when their family
moves to a new apartment. Other reasons for involuntary school changes are expulsions from
schools or the closure of the school that is currently attended. Students who are living in low-
budget housing, who get expelled from schools and whose school is closed down, are all
likely to score below average in tests. Comparing school-changers to non-changers would
thus partly represent differences of the students prior to the school-change.

To isolate the effects of changing schools on students, it is necessary to carefully choose the
control group. The most promising approach I have found in the literature focuses on students
who expressed the wish to change schools, but of whom only a random selection was allowed
to do so. Those students who were randomly denied the change are a perfect control group to
show the effects of switching. To my knowledge, there are well-made studies about three
choice systems in which such a control group can be found. In each of these, choice was
universal, applications to schools were rather easy to do and a considerable share of students
successfully changed schools. Moreover, admission to schools was influenced neither by
preferences of schools nor was it predetermined by the home location of the students. Thus,
all the potential problems for empirical research that were described in this section did not
apply to these three studies.

The first example is the introduction of school choice in Tel Aviv. Each student could choose
among 5 schools. Students were then assigned to one school of their choice. The assignment
was based only on a lottery after 2003. Before 2003, the aim of the school authority to
achieve a desired socio-economic and ethnic mix had an influence on student assignment.
Victor Lavy used the fact that students who did not choose their neighbourhood school had to
take buses to get to school. The author used the distance to the next stop of a school bus as an
instrument for the probability to apply to other schools than the neighbourhood school.

13 In each case, the home location had some influence. In Tel Aviv, the five schools among which each student
could choose were determined based on the home location. But students were admitted with priority only at their
neighbourhood school (Lavy 2006). In Charlotte-Mecklenburg school district, only students setting the assigned
home school as first choice had priority. But only about 50% did so (Hastings et al. 2005). In Chicago, students
were assigned a neighbourhood school where they were guaranteed a place. But about 50% attended other
schools than the assigned one (Cullen et al. 2003). Both in Chicago and Charlotte-Mecklenburg, the distribution
of places among students who were not assigned to the school was random. Thus, in all three cases, admission
into any school but the neighbourhood school was random and independent of home location.

14 Self-selection into the treatment is a potential problem. All students could choose schools, but those students
who actually exercise their right to choose, and are willing to accept a longer commute, are more likely to be
academically successful. Therefore, Lavy instrumented for the probability to choose schools by using the
Using this approach, he could identify a good control group. The author did not find higher gains for school-changers than for non-changers. However, compared to a good control group in neighbouring districts that could not choose schools, he did find sizeable gains for all students as a result of the competitive pressure faced by schools (Lavy 2006).

The second example is the Chicago Public Schools. There, all students can apply freely and admission is based on lotteries at most of the schools. By comparing lottery winners to lottery losers, three studies (Cullen et al. 2006, Cullen et al. 2007, Cullen/Jacob 2007) were able to identify suitable control groups. As admission was random, there should theoretically be no systematic differences between the winners of the admission-lottery at a school and the students who had applied to the same school but did not win. And all three studies confirmed that the treatment group and the control group were indeed not distinguishable according to the extensive available data. The students who won lotteries\textsuperscript{15} got into schools that had scored better on tests, had higher graduation rates and lower drop-out rates. And these schools had a better student intake, both according to socio-economic status and according to prior academic achievements. These schools were also located in areas with less crime and a higher average income (Cullen et al. 2003). But the lottery winners did not gain academically. Both the winners of entrance lotteries for elementary schools and for high schools did not score better on centralised tests than the lottery losers. Not even after several years at the seemingly superior schools, when the impact of a potentially better learning environment would have had the time to develop its full effect.

This result is similar to the one found by Lavy. And it is puzzling. The new schools seemed to be significantly better than the neighbourhood schools that the school changers left, according to most if not all publicly available data. But the students did not gain academically from the change when compared to their peers who did not win the lottery and had to stay at their neighbourhood schools.

One possible explanation is that the new schools were not really better at educating students. I have already hinted at the effect of different levels in the quality of student intake. It is distance to the next bus stop. This distance has an effect on the commuting time and thus on the probability to choose a distant school. And it is independent of innate ability. It is thus a viable instrument.\textsuperscript{15} Cullen et al. compared lottery winners to lottery losers. They did not compare students who actually changed schools to all those who did not change or to the lottery losers. This approach is necessary to deal with the problem of self-selection into the treatment. But it also dilutes the strength of the effects of switching. Only 20-35\% of the winners of each lottery took the place they had won. As many students applied to more than one school, some of the winners who did not take the place have probably won a lottery at another school and have taken the place there. And some of the losers might have also won another lottery and taken a place at the respective school. Comparing lottery winners to lottery losers while ignoring their following actions will thus show effects that are of a smaller magnitude than the effect that is to be expected as a result of an active school choice on an individual student.
possible that the popular schools are not really good at teaching students, but only scored well in the centralised tests because the local student intake is good. The authors of the studies that used data from Chicago suspected the same and checked value-added measures of school quality. It turned out that the schools that were identified to provide high value-added were slightly less popular with parents. Take-up rates of places at these schools after winning the lottery were slightly lower than for other schools and slightly more students of the high value-added schools applied to other schools (Cullen/Jacob 2007).

Thus, students who changed schools did not attend better schools according to value-added measures. They went to schools with good peers and to those that had scored well in tests based on pure outcome measures. Information on results of such tests was easily available to all parents. Information about the educational productivity of schools on the other hand was not available. Thus it comes as no surprise that parental choice was driven by outcome measures.

The third example is school choice in the Charlotte-Mecklenburg school district. There, students could hand in a list with three ordered school choices. Assignment was based only on these prioritised preferences of the students and on the capacity of schools. Again, students who got randomly admitted to seemingly good and therefore popular schools did not gain academically on average. But on closer inspection, the students whose parents had shown in their choices that they care strongly about academic quality did gain academically. On the other hand, students whose parents did not care strongly about academic quality actually lost academic quality as a result of the school change (Hastings et al. 2006). Those parents who cared more about academic quality were less likely to be from a minority and more likely to have a high-SES background. The authors analyzed in another paper what had driven the differences in the preferences for academic quality. One part of the explanation seems to be location. Low income and minority parents often lived in areas in which most schools had not scored well in previous years. For a significant increase in average test scores over that of their neighbourhood school, they would therefore have had to accept a larger increase in commuting time than high-income and non-minority parents.

Information also had an influence on school choice. When some parents were given additional information on the test results of available schools, the average test score of the chosen schools increased strongly (Hastings et al. 2007). This indicates that parents from low-SES backgrounds had trouble to identify the high schools that had a high quality, even according to simple average test scores that were publicly available. Once given information in an easily usable way, these parents made use of it.
3.4.2.(3) Effects of commuting to school

Proximity of the school, and thus a short commute to reach it, is one of the major concerns of parents. Proximity has been found to be an important issue for parents when making the school choice decision, both according to statements (Fossey 1994, Woods 1996) and actions (Schneider/Buckley 2002). Moreover, the distance to a school has been shown to limit the pursuit of academic quality when making the school choice decision. Parents have stated in interviews that transportation concerns deter them from choosing academically better schools that are farther away (Goldring/Hausman 1999). In Chicago, parents tried to avoid choosing schools if their children would have to cross gang territory to reach them (Lauen 2007a).

Students in Chicago were also much more likely to opt out of an otherwise similar school if there were more alternative schools close by, so that a school change would mean a lower increase in commuting time (Cullen et al. 2000). In the Charlotte-Mecklenburg school district, another mile distance reduced the odds of choosing a school by 27-35% (Hastings et al. 2005).

To sum up the presented evidence: parents care much about proximity of schools and their school choice decisions are strongly influenced by the distance to potential schools. Parents explicitly trade off academic quality for proximity when choosing schools.

Thus it is not surprising, that if students opt out of neighbourhood schools, the commuting distance does not increase by much. In the Charlotte-Mecklenburg school district, the estimated odds described above mean, that a school that was three miles farther away would have to score an average 1-2 standard deviations better on the available quality measures to be chosen (Hastings et al. 2005). This does not mean though, that individual students did not travel much farther. There were two groups of students in the district with diverging choice behaviour. One group essentially just went to the closest school, regardless of academic quality. The other group shopped for schools with high perceived academic quality and was willing to accept a long commute (Hastings et al. 2005). The estimates given above are an average across all students.

In Chicago, lottery winners travelled only an average of half a mile farther than their peers who had lost the lottery for entrance into a popular school (Cullen et al. 2003). This figure likely understates the real distance though, as only about a third of the lottery winners actually took up the place they had won and it is not possible to take the outcomes of other lotteries into account. But even three or four times the observed half mile does not lead to a largely increased commuting time. But again, this is only an average. Individual students commuted much farther. It is also worth to mention, that the density of schools in the inner city of
Chicago is rather high, so that there are several schools within a 1.5-2 mile distance to the home of most students.

Summing up, there are costs connected with a longer commuting distance that parents and students want to avoid. These costs might also have an effect on student outcomes. It would be interesting to know how strong this effect is. But it is probably impossible to disentangle the effect of commuting costs from other effects that result from school changes. I will deal with those effects summarily after part (5) of this section, once I have presented them all.

3.4.2.(4) Adverse effects of changing schools

It is possible that changing the school has a detrimental effect on academic outcomes. The student has to get used to new teachers and peers. She will likely be separated from some of her friends and classmates whom she has known, and was used to, from previous school years. Commuting to a non-neighborhood school also probably means, that there are fewer classmates close by to learn together with or to bring lecture material when sick. Teaching methods and curriculum might change. It is possible that some content was already covered in previous grades in the new school but not in her old school, so that she has to catch up. As in part (3), these possible effects of changing schools are tightly entangled with other effects and will be addressed after part (5).

3.4.2.(5) School changes and the match between school characteristics and preferences

It should be possible to see the effects of a better match between preferences of parents and students and the school. If the match improves, satisfaction should increase. The introduction of school choice usually does lead to higher levels of satisfaction with the school, both for parents and for students (see for example Lavy (2006)). This increase in satisfaction is probably induced partially by a better match between preferences and the curriculum, education style or other features of the school, and partially by an increase in perceived quality\(^{16}\). This would explain why satisfaction with the school increases for students who actually change schools. But surveys taken after the introduction of school choice often show improvements in the satisfaction with school for most parents, not only those of students who have changed the school. In Tel Aviv for example, satisfaction with schools increased after the introduction of school choice. And there was no significant difference in the size of this

---

\(^{16}\) In most school choice systems, parents who value academic quality send their children to schools which provide good peers and that scored well on standardized tests, as these statistics are usually the only available academic outcomes of schools. As presented in part (1) of this section, these schools are often not really better at educating students. But as long as parents think that their child got access to a better school, they are likely to be more content with it. Usually, these questions are asked shortly after the school change, so that there is no time for parents to realize that the grades of their children do not improve much after the change.
increase between the students who had changed schools and those who had stayed in their neighbourhood school (Lavy 2006).

There are two possible sources for the increase in satisfaction of non-changing students. The first is that schools try harder to satisfy students and parents so that these do not leave the school. This can affect the satisfaction of all students. The second possible source is the possibility to choose itself. A separation of these two sources is probably not feasible. Only if parents have a real possibility to choose should they be able to draw satisfaction from this fact. But if they have a real possibility to choose, then schools have to fear losing them and probably try harder to satisfy students and parents.

I still have to discuss the strength of three effects of changing schools on academic outcomes: of a better match between student needs or preferences and school characteristics, of detrimental effects of changing schools and of a longer commute. How do these separate effects add up in affecting academic achievement? This question is very hard to answer, maybe even impossible to answer conclusively. The reason is that several parameters change simultaneously when a student changes schools. The student usually has a longer commute to school. She also changes to a new school which might be a better match for her needs which in turn could affect academic success directly. The new school might also be a better match for her preferences, which should improve satisfaction and motivation and thus affect academic success indirectly. And the student has to adapt to new peers, teachers and a new environment. Additionally to these three effects of immediate interest, the school might also be better or worse at educating students than the old one. So far, I have not seen a study that successfully disentangled these effects.

Therefore I have to resort to informed guessing: I think that it is likely that a long commute ceteris paribus harms academic success as it is tiring and costs time. I suspect that being separated from old friends is on average slightly detrimental to satisfaction with the school and therefore for academic outcomes. If those old friends represent what is called a “bad influence” however, that separation might be beneficial for academic achievement. Examples would be students who opt out of neighbourhood schools in areas with a low SES and a strong influence of gangs on adolescents. Finally, I suspect that a better match between school characteristics and preferences increases satisfaction. But I only expect increases in academic achievement from a better match if parents are able to spot characteristics of individual schools that are academically beneficial for their child. Then they have to give these characteristics a high priority when choosing a school. Given the evidence presented on the
quality of information on educational quality, I doubt that a large share of parents succeeds in identifying the schools where their child could learn most.

Fortunately, for the mechanism of Friedman to work, it is only necessary that the aggregate effect of choosing a school with a higher academic quality is positive on average. Therefore, the relative strength of individual effects is not important and I do not have to rely on my educated guesses.

3.4.3. Summary on students and the effects of changing schools

Choices about schools are dominated by preferences of the parents. This might be because the preferences of students are similar to those of the parents or because parents have the final word. But in either case it is not necessary to find out about the preferences of students. The preferences of parents can be used to explain school choices made by households.

Besides the academic quality of schools, parents care most about the proximity of schools to the household. The average commuting distance to chosen schools is therefore not long on average.

The effects on academic achievement of the commuting distance, a changed environment and a potentially improved match between student needs and school characteristics have so far not been disentangled. Nor have these effects been separated from variations in the educational productivity of schools. For school choice to improve overall educational productivity via competition, the change to a popular school has to result on average in an increase in academic achievement. For this to happen, it is only necessary for the aggregate of these four effects to be an increase in academic achievement. Thus, it is fortunately not necessary to disentangle the effects for the purpose of my dissertation.

At the first glance, previous studies had difficulty to identify such academic gains as a result of changes to seemingly better schools. On closer inspection however, the better outcomes of the popular schools owe more to a better student intake than to a higher educational productivity. Thus, a school change to a “better” school did on average not result in the student receiving higher quality education. Additionally, students whose parents cared more about academic quality, or whose parents were better informed, do show some academic gains.
3.5. Information

All decisions that are taken by actors in school choice systems are based directly or indirectly on information about the characteristics of schools. Parental school choice is based on academic quality measures and information about other characteristics of schools. The effort decisions of individual schools and teachers are based on the expected loss of students and consequences associated with this loss. The share of students that an individual school expects to loose depends in turn on the quality of competing schools. Finally, any sanctions and benefits for schools by school authorities are usually based on information about individual school quality or on the number of lost students (which in turn depends on school quality).

Actors gain information about school quality from formal and informal sources. Examples for informal information are second-hand experiences from friends and relatives, the reputation of a school and impressions from visits to school campuses. Formal information sources are for example test outcomes, class size and summary statistics about schools like graduation rates, truancy rates and the composition of the student body. Such school characteristics are often published in newspapers, on the internet or handed out to parents on school report cards.

While school authorities can hardly affect the informal information about schools, the formal information is at their discretion. School authorities can choose whether to publish any information at all. If they choose to publish, they can choose about which characteristics information is made available and which distribution methods are used. School characteristics could for example be handed out only to current students, to all potential students or posted freely available on the internet. Moreover, school authorities can decide on how to generate the statistics that will be published.

There is not much leeway in how to define the share of male students or the ethnic composition of a school. For most summary statistics, the decision is therefore simply whether they will be published or not and which medium will be used for publication. For information on academic quality however, school authorities can choose among several ways to measure individual academic achievements and how to aggregate these into school characteristics.

3.5.1. Different possibilities of measuring and presenting academic school quality

There is a wide variety of measures that are used to represent the academic quality of individual schools. To structure the presentation of these measures, I have identified the four most important dimensions on which the individual measures differ. These are: (1) the
method of measuring student outcomes, (2) whether individual student characteristics were taken into account, (3) the methods used for aggregation and/or representation and (4) whether tests are administered centrally or not.

3.5.1.(1) The methods for measuring student outcomes

There are three main methods that are used to measure academic achievement of individual students: raw test outcomes, gains and year to year changes.

The first method uses the test outcomes that individual students scored in tests in one year. Measures of this type do not use any information of outcomes in previous years, neither of the student nor of the school. The outcomes of one test only are used to compute this type of quality measure, which makes it rather easy to generate. But quality measures that are based only on the outcomes of one test do not reflect the educational quality of the attended school directly. Such measures are also affected by differences in average student characteristics (see part (2) of this section below). Moreover, raw outcomes also reflect the effects of education in previous schools. With outcomes from only one year, it is not possible to determine which share of the measured achievement of a student is attributable to the current school and which share is due to the student’s previous education.

The second method uses gains of students between two points in time. This method analyzes what individual students have learned between two tests. To implement this method, it is necessary to track students over time and to match the results in at least two tests for each student. If a student changes schools between the two tests, the interpretation of gains becomes complicated. Gains could, for example, be due to a good quality of education at the second, or to a bad quality of education at the first school. Moreover, students who have achieved good results in previous years might be in a better position to learn more, as they command a sound understanding of basic knowledge. On the other hand, it might be easier to catch up from a low level than to make further gains on top of good results in the first test. The net impact of these two effects is not clear and further complicates the interpretation of differences in gains.

The third method that is commonly used to measure student achievements, uses year to year changes. This method compares the outcomes that the students in a given grade of an individual school scored, to the outcomes scored in a previous year by another cohort of students who then attended this grade. Measures generated with this method are not a convenient way to compare academic quality across schools at one point in time, unless all schools start out at the same level. Instead they show the progress that was made by the
students of an individual school over time. Moreover, variations in this measure are driven not only by changes in the quality of education at individual schools but also by variations in the quality of the yearly student intake. Especially small schools are strongly susceptible to changes in the student intake. A few exceptionally bright or disruptive students can induce significant variations in yearly progress for a given level of educational quality. If there are a few exceptionally bright students in the intake of one year only, this spike in student quality alone would for example translate into a strong apparent increase in quality in the year in which this cohort is tested and to a strong decrease in the following year. Another drawback of year to year changes is, that such measures can be distorted by trends that are not connected to changes in educational school productivity. If, for example, crime rates increase strongly in a neighbourhood over several years, most households that can afford to do so move away within a few years. Those who stay are likely to have a rather low income and therefore no choice but to stay. Their children usually do worse at school than those of more prosperous parents. The gradual change in the student population, which is induced by households that are moving away, would lead to several consecutive year to year losses in scores for schools in this neighbourhood. Successful earlier reforms in lower grades can also distort measures of year-to-year gains. If students who profited from such reforms replace those who did not, this would result in an apparent progress of the school, again without any change in the quality of the education (Meyer 1997).

3.5.1.(2) Accounting for individual student characteristics

Students are not identical at the beginning of a school year. Some have more previous knowledge than others, some obtain more support from their parents than others, and some have been socialized in an environment where almost every adult they know attained higher education and works in professions that require such an education. Such students are likely to be more motivated at school than those who mostly know adults who are employed in low-skill jobs or unemployed. Moreover, some students do not have a full command of the language used in school or have disabilities that slow down their learning progress. Differences in the students’ family background are likely to affect test results if they are systematic. Most schools cater mainly to the local student population, and household location is partially driven by characteristics that also affect the family background of students. In low-rent areas live, for example, mostly adults who are working in low-wage jobs. These adults have, on average, a low education and are therefore not in a good position to support the
education of their children. Due to such mechanisms, differences in average family background across schools are likely to be systematic.

And differences in family background affect test outcomes. If systematic differences in family background are not taken into account, they will therefore distort differences in test results. This means that it is hard to find out how much of the difference in test results between two schools is attributable to variations in school quality, and how much is attributable to differences in the characteristics of the student intake. As factors that are unfavourable for good test results tend to cluster in low-income areas, differences in the student intake can potentially drive most of the variation in test results across schools. If parents base their school choice decisions on measures that are not driven by educational school productivity, this might have important consequences for the outcomes of school choice. Schools that have a good student intake can document good quality measures if these are based on raw test outcomes. These good quality measures then attract many applications. Good students are more likely to apply to schools than academically weak students (see 3.1.2.(4)). Schools that at one point of time are able to document favourable test results are therefore likely to receive an above average student intake. This mechanism can perpetuate effects of a good student intake: a school has a good student intake which results in good test results which in turn attracts a good intake.

But if variations in student intake are taken into account, it is possible to control for the effects of differences in student characteristics. Then, it is possible to isolate the educational value-added, or educational productivity, of schools. In order to identify this educational school productivity, it is necessary to capture most student characteristics that have a significant effect on test results. Among the student characteristics that were identified to strongly affect test results are gender, within-grade age, previous academic achievements, parental education and migration background (Wößmann 2003). Limited proficiency of the local language, disabilities and summary characteristics of some of the above for the peers also affected test results. In the absence of detailed information on parental background (such as education, income or support given to children in learning), race and low income status are usually highly significant proxies.

Value-added quality measures take differences in student characteristics into account. To generate value-added quality measures, the first step is to regress individual test results on student- and school-characteristics. The results of this regression are then used to predict expected test results for each student, given her characteristics. This procedure yields, for example, the predicted test result for a Hispanic girl who is young for her grade, who receives
lunch at reduced prices, has no disabilities and no difficulties speaking English and scored well in the test last year. In a third step, value-added outcomes are calculated for individual students by comparing the predicted test results to the results that were really scored in tests. The difference between these two results then shows, how the test result of an individual student deviates from the average result of students who share her characteristics. This difference should represent the effects of the quality of education at the attended school. It is therefore called the value-added that is provided by the school.

3.5.1.(3) Aggregation methods
There are two common methods to aggregate the test results of individual students. The first method is to simply calculate the average score of all students at a school. With this method, test results of all students count equally. For schools that intend to improve their test results, it is thus most effective to focus on those students, who are likely to make the most progress for a given amount of teacher effort.

The second common aggregation method is to report the percentage of students at a school that managed to perform above given thresholds. Such thresholds might be the average or median, national test result or a fixed number of points scored in a test. In this case it is most effective for schools that intend to improve their test results to focus on those students who are likely to fail the threshold by a narrow margin. Effort spent on students who will fail or meet the threshold irrespective of teacher effort will not have an effect on the measured school quality if this method is used.

Both types of aggregation can be represented for the entire student population. Moreover, it is possible to break down these characteristics for different student groups. Aggregate test outcomes could be represented for boys and girls separately. They could also be broken down by migration status, ethnicity/race, previous academic achievement or low-income status of the parents. Such a break-down would provide a more accurate prediction to parents for test results to expect for their child at an individual school, given her characteristics. On the other hand, breaking down aggregate outcomes by too many characteristics is likely to create small groups of students. Estimations for these might be driven by a few exceptional students in the group and thus not be suitable to predict outcomes for other students who share the same observable characteristics.

3.5.1.(4) Centralized tests
School quality measures could be based on centralized tests or on tests developed and carried out by individual schools. If the design, content and grading of tests is not determined
centrally, the tests are likely to vary in difficulty and/or in the content that they cover. Thus, the outcomes of such tests cannot be compared easily. Moreover, if test results are published and it has consequences for the school how it scored in comparison to other schools, it is tempting to lower standards in order to achieve higher test results. Therefore, only central tests provide a sound basis for the generation of quality measures that can be used to compare schools.

3.5.2. Observations and empirical evidence on academic quality measures

Summary statistics of the characteristics of students are made publicly available in most school choice systems. Information on enrolment numbers, class size, gender and ethnic/racial composition are usually made available on homepages or school report cards. Information on drop-out, attendance and truancy rates or on the qualifications of teachers is, on the other hand, not available in all school choice systems. Generally, the extent of information that is available varies greatly. And at least in the case of school choice in Sweden, hardly any characteristics were made public in a centralized way (Söderström 2006). If summary statistics of the composition student body are given, parental choice behaviour reacts to this information (see 3.1.2.).

Some kind of academic quality measure is publicly available in all school choice systems that I have mentioned so far, again with the exception of Sweden. But the measures that were used vary greatly on most of the dimensions presented in the previous section.

3.5.2.(1) Methods for measuring and presenting student outcomes

Quality measures based on raw test results are the most common in reality. Among the school choice systems presented in this text, they are used in England, Chicago and Charlotte-Mecklenburg County. Several accountability systems like the No Child Left Behind (NCLB) policy also use such measures to determine whether a school is failing.

Gains of individual students are used less often. If so, they usually complement existing measures based on raw outputs. That these measures usually appear in tandem is no surprise, as the information that is necessary for raw output measures is a prerequisite for the generation of measures based on student gains. A typical example can be found in the CPS. There, the prominent measures for the academic quality of schools are based on raw outcomes of tests like the Prairie State Achievement Examination (PSAE) or the American College Test (ACT). These are the measures that are reported on school report cards, the medium that
parents most likely use. But on the homepage of the CPS, one can also find gains measures for the progress of individual students between two such tests (CPS School Data).

Year-to-year gains for schools are mostly used to evaluate the progress of schools over time. The NCLB policy for example uses yearly progress to evaluate whether schools that are on probation are on the right track. If a school that already is on probation fails to meet the required progress, students are allowed to leave the school. The accountability programs of US-states such as California or North Carolina also use year-to-year changes to assess schools. As mentioned above, year-to-year changes are easily distorted by changes in the student intake. A few exceptionally bright or disruptive students can have a strong effect on changes. The smaller the school is, the stronger measures of change will be affected by a few outlier students. Consequently, virtually all schools with exceptionally high or low gains in North Carolina had low enrolment numbers. And there was a strong mean reversion, meaning that schools with exceptional average test results in one year had average results much closer to the state average in the following year (Kane/Staiger 2002). Expressed in year-to-year gains and losses this means, that an exceptionally high gain in quality was usually followed in the next year by a considerable loss. The correlation for changes in fourth-grade math between one year and the next was for example estimated at -0.37. This is a strong indication that the year-to-year changes did not capture differences in academic quality. If gains are due to an increase in educational productivity, they should persist and students in the following years should score as high as the previous cohort or even higher. If, on the other hand, gains are due to a few bright students in the intake of one year, there is no reason to expect any effect on following cohorts. Estimations put the share of year-to-year changes in North Carolina that was non-persistent at 70-90% (Kane/Staiger 2002).

3.5.2.(2) Accounting for individual student characteristics

Of the variance in test results of students in North Carolina, only about 10-15% is due to variance across schools. The remainder, by far the bigger share, is within-school variance across students (Kane/Staiger 2002). These estimates mean that there exist strong differences between test results of individual students who are educated in the same school. It also means that differences in student characteristics have a stronger effect on individual student test results than differences in the academic quality of the attended schools. If the characteristics of the student intake differ in any systematic way, average school test results are therefore likely to be strongly driven by differences in the student intake.
Student characteristics do vary systematically across neighbourhoods and most schools are mainly attended by students who live close by. Therefore, student characteristics do vary systematically across schools. What it means for the precision of school quality measures if they do not account for systematic differences in the student intake can be learned from an exercise done by Wilson and Piebalga (2008) on UK data. The authors generated educational school productivity measures that controlled for student intake and compared those to measures based on raw output that did not take student characteristics into account. They found a correlation between educational school productivity and measures based on raw output of only 0.33 to 0.46, indicating that differences in average test results were also strongly driven by variations in the student intake (Wilson/Piebalga 2008). These correlations were computed for all secondary schools in England. In areas where the variation of social status across neighbourhoods is more pronounced, as in the case of big cities, the variation in student characteristics is likely to be higher and therefore likely to have an even stronger effect on average test results. Douglas Lauen found for such an area, the school district that covers the inner city of Chicago, a correlation between educational school productivity and test scores that was slightly negative (Lauen 2007a).

More evidence on the effects of characteristics of the student intake on test results can be gained by analyzing the intake of schools that do well according to measures that do not account for differences in the student intake. In the UK, schools that are deemed to be good, based on the most prominent output-based measure, are most likely to have a socially advantaged student intake and are most likely to be academically selective or to have been academically selective before (Whitty 1997). Selective schools can pick students whom they expect to perform well in test. Prosperous areas tend to be populated by households with highly educated parents who speak the local language fluently and to whom academic success of their children is important, characteristics that are favourable to learning. And rules in England favour the admission of local residents into schools, meaning that sought-after schools are almost exclusively populated with local students (Burgess et al. 2004, Whitty 1997).

Academic quality is one of the strongest drivers of parental school choice (see 3.1.2.(1) and (2)). And to gauge the academic quality of schools, most parents rely on official quality measures which are usually based on raw test outcomes. The evidence presented above implies that school quality measures which are based on raw test outcomes are to a large extent driven by differences in student intake. Thus, when using those official quality
measures, parents do not identify the highly productive schools as “good” but the ones with a good student intake.

One example for the use of purely outcome-based quality measures is Chicago. Schools that had previously scored high average test scores received more applications. But by examining educational productivity measures, the authors found that among those schools that had received any applications, school popularity decreased with educational productivity (Cullen/Jacob 2007).

This inability of parents to identify highly productive schools can break the mechanism that was described by Milton Friedman: if parents cannot identify the more productive schools, these do not attract more students and are thus not rewarded for their higher educational productivity. Without rewards, there is no reason to expect that schools try to increase educational productivity as a reaction to competitive pressure.

Instead, schools will try to improve their performance according to the quality measures that can be observed by their customers, the parents. Thus, in a school system that is otherwise favourable to effective school choice, competitive pressure based on raw outcome measures will induce schools to increase outcomes, not educational productivity. Although a school can increase outcomes by increasing educational productivity, there are other means that require less effort and can also achieve the same goal. Therefore, with quality measures based on raw outcomes, competitive pressure might not give incentives to increase educational productivity, but give incentives to divert efforts from teaching into cream-skimming, finding ways to classify weak students as disabled (Wilson/Piebalga 2008), outright cheating or other detrimental activities as described in 3.2.2.(3). Kane and Staiger (2002) have expressed this insight of the dependence of incentives on the type of quality measures nicely: “Many accountability systems that appear reasonable at first glance perform in perverse ways when test score measures are imprecise.” (Kane/Staiger 2002 p.91).

For the same reasons, information has been named the Achilles Heel of school choice (Teske/Schneider 2001), implying that a flaw in the parental information on school quality could bring down the entire mechanism that is expected to improve educational school productivity by means of school choice.

17 Kane and Staiger comment on accountability systems in general. These include direct rewards and punishments of schools. Although the punishments include the right for students to leave failing schools, as according to the NCLB policy and in some voucher schemes, accountability systems are not the same as indirect pressure from school choice. But in both cases incentives are created based on generated quality measures. Therefore, this comment also applies to the situation in school choice systems.
The fact that parents confuse the fact that a school achieves high test scores with the fact that it is more productive might also explain why it is so hard to show gains for students who opted into seemingly better schools, as discussed in 3.4.2.(2).

Researchers and school authorities have become aware of the problems associated with using raw test outcomes. Among the studies mentioned so far, Cullen and Jacob (2007) computed value-added measures to estimate the educational productivity of schools in Chicago. Wilson and Piebalga (2008) did the same for schools in England and Jacob/Lefgren (2005) for teachers in a mid-sized school district in the US. Robert H. Meyer of the University of Chicago has also worked on value-added measures for schools, for example in Meyer (1997).

In England, a value-added measure has been included in the league-tables since 2002. This measure only uses previous test results of students though, and does not account for differences in student characteristics such as gender, age and ethnicity. And it is largely ignored by parents (Wilson 2003). Another measure, that accounts for a multitude of background-characteristics, was also introduced in England, but only in 2006 so that it has been hardly exploited for research so far, with the exception of the paper by Wilson and Piebalga (2008). The CPS has also recently introduced a more thoroughly constructed value-added measure in cooperation with R. H. Meyer. This measure represents gains of students within one year and accounts for student characteristics such as an English language learner status (CPS Value Added). This measure was first published in early 2009 for gains of elementary school students between 2007 and 2008.

3.5.2.(3) Aggregation method

All three types of outcomes (output measures, gains and year-to-year changes) are presented in diverse ways. Measures based on raw outcomes or gains are usually presented either as averages for schools or as the percentage of students who managed to exceed given thresholds. These thresholds usually are test point scores for raw outcome measures and a required yearly progress for the gains measure. Both ways to present the data are easily generated from the same data source. It is thus common for school authorities to make both available. But usually one is given a prominent place in the publications and parents tend to focus their attention on only one measure. In the UK for example, newspapers publish a wide selection of different quality measures. But parents base their decisions almost exclusively on the “%A*-C”, a measure that gives the percentage of students who passed classes in at least five subjects with at least the grade C (Wilson 2003).
Measures based on year-to-year changes often categorize schools into those making “adequate progress” and those failing to do so. This is for example the case for the NCLB policy. This categorization creates an artificial dichotomy so that a small difference in average test results determines whether a school is signalled as failing.

For all three types of measures it is also common to create a ranking of schools according to the most prominent measure. Such a ranking might be provided by school authorities or by newspapers that use data presented by the school authorities in different ways. One example are the league tables in the UK that list all secondary schools in nation-wide rankings based on the %A*-C measure\textsuperscript{18}.

All three types of outcome measure can also be aggregated either for all students, or they can be stated for sub-groups of students separately. Students are sorted into subgroups for example by ethnicity, whether they have sufficient command of the local language, by low-income status and by gender. The results of the PSAE test in the CPS can for example be accessed on the internet separated by race, gender, limited English proficiency status, low income status and individualized education program status. On the CEO school report cards of the CPS, the source of information on schools that is most easily accessible to parents, only school averages are reported however. The rankings based on English league tables sort schools using only aggregate results for the entire student body (DCSF 2009).

\subsection*{3.5.2. (4) Centralized tests}

Tests that are not administered centrally do not allow a meaningful comparison across schools and, moreover, set incentives for grade inflation. They are only used in one school choice setting that I am aware of, in Sweden. However, to my knowledge there was no information on school outcomes that was published centrally in Sweden.

\subsection*{3.5.3. Summary on information}

Some information about summary characteristics of schools, such as the share of female students or graduation rates, is publicly available in almost all school choice settings. While ways to present information on such characteristics are limited, there is a multitude of ways in which to present measures on academic quality.

The most common measures on academic quality are based on raw test outcomes, usually aggregated by taking school-averages or by presenting the share of students that managed to

\textsuperscript{18} In case of a tie in the %A*-C measure, average grades are used to determine a ranking order.
score above given thresholds. Gains and year-to-year changes, the other popular measures, are harder to interpret when used for comparisons across schools and are less common.

In most school choice systems, all three types of academic quality measures do not take differences in the characteristics of the student intake into account. However, average student characteristics strongly affect test results and often differ systematically across schools. If such systematic differences are ignored, they distort academic quality measures, heavily favouring schools whose local student population is academically advantaged.

A consequence of such a distortion of quality measures is, that parents who care about academic quality cannot identify the highly productive schools. With quality measures based on raw outcome measures, parents favour schools that show good outcomes, which are mainly due to characteristics of the student intake. Thus, schools are not rewarded for high educational productivity but for an academically advantaged student intake. In this case, schools do not face incentives to increase educational productivity but face incentives to improve their intake through selective measures.

A potential remedy to this distortion of quality measures is to use measures that account for differences in the student intake and thus isolate the effects of educational school productivity. Some research has already been conducted on such value-added measures. Implementation however has so far been limited to an imperfect and largely ignored measure in England, and two proper measures, one for elementary schools in the CPS and one for secondary schools in England. Both proper measures have just recently been published for the first time. The second part of this dissertation (starting in Chapter 5) aims at exploring the potential effects of a prolonged use of value-added measures in school choice systems.
4. Approaches to analyzing the effects of school choice in previous literature

The empirical literature on school choice that I have presented in the last chapter has produced conflicting evidence in several areas. Moreover, this literature has so far not been able to validate several mechanisms that many researchers expected to find. Take, as an illustrative example, the effect on students who were admitted at seemingly better schools. Most researchers expect that schools differ somehow in their ability to educate students. But it is hard to find academic gains for students who change schools. Even if the new school seems to be superior in almost all respects that are deemed important for academic quality. Most studies did not find any effects of changing schools on academic achievement (see for example Cullen et al. 2003), while some found at least limited effects for subgroups of students (for example Hastings et al. 2006). But no study found effects of a size that could be expected as the result of a considerable improvement in school quality.

The conflicting evidence, and the inability to document intuitively convincing effects, are probably due to several characteristics of school choice that complicate research in this field. In this chapter I will present these characteristics and show how they have affected previous research. I will start with general problems that affect all approaches to investigating school choice (4.1). Then I will focus in turn on empirical (4.2) and theoretical (4.3) approaches that have been used to analyze school choice. In each section I will present additional problems associated with the specific approaches discussed. In each section I will also describe some studies that have used the respective approach, how these studies tried to cope with the presented problems and which insights the authors were able to gain. I will conclude this chapter with a summary of implications and on how to best deal with the identified problems (4.4).

4.1. General difficulties

4.1.1. Differences in institutional settings and characteristics

School choice takes place in a highly regulated setting. Rules that are set by school authorities determine, for example, which students are allowed to choose schools, how schools are to select among applicants, which information is made available to parents and what happens to schools that loose students. As demonstrated above, several of these regulations can break the
mechanism that Milton Friedman described if they are set in a specific way. Sometimes, a specific combination of settings in several regulations can break the mechanism (see 3.3.1.(7) for two examples). And certain settings in some regulations can break the mechanism on their own, for example if only a small share of students is allowed to change schools, or if no financing follows students. Moreover, schools might try to game the system when the loss of students has unpleasant consequences and there are loopholes in the regulations that can be exploited (see 3.2.2.(3)).

Thus, it is necessary to analyze all regulations and settings of a school choice system carefully, before it is possible to gain insights that might apply to other school systems than the currently analyzed one. But these settings are often not easy to identify for a given school system, and the actual effect of several regulations is not obvious at the first glance. As schools might try to circumvent regulations, it is necessary not to rely on official regulations, but also to analyze the actual behaviour of schools and students in a choice system. Summing up, a quite intimate understanding of the regulations and mechanisms of the analyzed school choice system is necessary in order to avoid pitfalls that might heavily distort any findings.

4.1.2. Interdependence of decisions

The second characteristic that complicates all research on school choice systems is the interdependence of decisions among actors. Every school and household decides for themselves, but all these decisions influence other actors, either directly or by changing the information on which the decisions of others are based. Here is one example that illustrates this web of interdependencies: if parents choose to apply to a popular school, they directly reduce the probability of all other applicants to be admitted. If the application is successful, they also reduce the number of students at the school that their child attended before. Additionally, the changing student will change summary characteristics and average test results at her old school and at the one which she now attends. Let us assume that the student is very bright. If other parents care about test results, the new school will then become slightly more attractive due to the good test result achieved by the new student. Additionally, the quality of other nearby schools will look slightly worse, relative to the new school of this bright student. These schools might decide to react to the increased competitive pressure by increasing effort.

These interdependencies make it hard to isolate an individual effect in empirical research and to catch its impact in full. For a theoretical approach, the described interdependencies give rise to even more serious problems (see 4.3.1.(1)).
4.1.3. Heterogeneous actors

The third difficulty that renders insights into the mechanisms that are induced by school choice difficult, is caused by heterogeneous characteristics and heterogeneous behaviour of actors. I have presented several examples for the heterogeneity of actors in detail above. The propensity of students to apply for schools is, for example, positively correlated with factors that also affect innate ability. And the willingness to accept a longer commute to school in exchange for a higher academic quality is more pronounced in high-ability students. Moreover, these differences are strong. In Charlotte-Mecklenburg County for example, one group of parents basically sent their children simply to the closest school. Another group of parents was willing to accept large increases in commuting time in exchange for slight increases in perceived academic quality (Hastings et al. 2006). Moreover, the students who are more likely to change schools for increases in academic quality are more likely to be of a high innate ability and a favourable socio-economic background. This has strong implications for sorting by ability and SES which would be missed when heterogeneity of the actors is not accounted for.

Any empirical research that does not account for heterogeneous actors, via interaction effects or by looking at subgroups of students for example, can only find average effects. Likewise, any theoretical approach that assumes representative agents is only able to represent an average effect that misses these differences in choice behaviour.

Apart from different preferences, parents and students also show different characteristics. These differences have an influence on how their decisions impact on other actors. If proximity to schools matters to parents, then the household location has a strong effect on probabilities to apply to individual schools. And household locations of individual students are not determined by chance. The inhabitants of a neighbourhood tend to be similar concerning many characteristics, mainly due to levels of housing prices that vary across neighbourhoods and are thus reflected in the average income of households living there. Thus, the local student population of schools is systematically affected by the distribution of characteristics in the local population. Moreover, the effects of students who change schools depend on their characteristics. If, for example, a bright student changes to another school, this results in lower average test results at the old school and higher average test results at the new one. If a weak student changes to another school, the effects are reversed. Another important difference across actors in real world school choice systems is the information on which they can base their decisions. High-SES parents usually are better informed about the
academic quality of individual schools. This informational difference translates into different choice behaviour.

Any empirical research that does not account for the heterogeneity of actors will miss much of what is going on. And a theoretical approach that uses representative agents would not be able to adequately represent the mechanisms that drive the outcomes of choice in school systems.

4.2. Empirical approaches

In this section, I will analyze characteristics that result in problems for empirical approaches to analyze school choice. I will start by presenting these problems (4.2.1.) and then show which effects they had in previous studies, how the authors of these studies tried to deal with these problems and which insights they gained (4.2.2.).

4.2.1. Problems for empirical approaches

4.2.1.(1) Effects of the application and admission mechanism

The regulations regarding applications and admission to schools can have effects that distort empirical estimates. I have discussed these effects already in 3.4.2.(2). Therefore I present here only a short reminder of the mechanisms that can result in systematic differences in the student intake of individual schools. These include: self-selection into voluntary school choice, cream-skimming by schools, involuntary school changes that are not driven by school quality considerations and preferential admission of local residents.

4.2.1.(2) Additional reforms, trends and delayed effects

The introduction of school choice is a mayor reform of the school system. Such mayor reforms are often accompanied by other reforms (Whitty 1997, Hoxby 2000). If several reforms are implemented at the same time, it is hard to isolate the effects of school choice. Also, many of the changes that are expected as the outcome of school choice take time to materialize. The effects on students who change schools might be distorted in the first year by the fact that these students need to adapt to their new peers, teachers, teaching methods and facilities, before they can fully benefit from a more productive learning environment. Schools need time to experiment with - and to implement - new teaching methods that might help to increase educational productivity. Weeding out the syllabus in order to focus on vital content can also not be done instantly. Some changes might even need decades to materialize fully. If school choice results in a working environment that rewards effort for example, schools might
attract a different breed of teachers. But this new breed will only slowly replace the older teachers as those retire.

Effects might not only be delayed. Moreover, outcomes in the first few years after the introduction of school choice might be misleading. Some students are more likely to make use of school choice than others. In the first years after the introduction of choice it would be mainly these students, who tend to be of a higher innate ability than the average student, who would change schools. Effects in the first years, especially on sorting and outcomes of individual schools, would thus be mainly driven by those students who make use of choice first (Hoxby 2000).

Delayed effects have two important implications. First, it may take several years before the effects of school choice are fully observable. Second, during the time that is necessary for effects to materialize, other reforms might be implemented or general trends on achievement, that are not related to school choice, might have an effect. In both cases, it would be hard to isolate the effects of school choice.

4.2.1.(3) Reverse causality

Reverse causality is another problem for the validity of empirical estimates in the field of school choice. Choice is more likely to be introduced in school systems that are systematically different from those where school choice is not introduced. Teacher unions usually are stalwart opponents of school choice schemes. And these unions often are quite influential in education policy (Moe 2001). To challenge teacher unions needs determination and/or public support for the change, both of which are likely to be stronger if the school system obviously is in need of a change in the perception of parents and/or voters. School choice was for example first introduced in the CPS to reduce strong segregation by race and the school system has disproportionate minority and poverty rates leading to poor overall performance in comparison to national norms (Cullen/Jacob 2007). The Chicago Public Schools were even named as the “worst in the nation” by the then U.S. Secretary of Education in 1987 (Lauen 2007a). The NCLB policy introduces school choice only for students of schools that failed to meet standards two years in a row (CPS NCLB). Charter schools are more likely to be founded in areas where parents are particularly unhappy with the public schools. And several voucher schemes, financed both by public funds or private donations, were introduced on purpose to provide choice to poor students in badly performing schools. Examples are voucher schemes in Milwaukee (Wisconsin), Washington D.C. and the area affected by Hurricane Katrina. Thus, school choice is more likely to be introduced in areas
with a weak academic performance. A cross-section estimation that compares outcomes of areas with school choice to areas without choice might therefore be biased downward when measuring effects of school choice on educational productivity. Moreover, it is conceivable that teacher unions are weaker in areas where school choice is implemented. And the fact that teacher unions are weak might have other effects on academic performance.

4.2.1.(4) External choice options

Options to choose schools outside the local public school system might also lead to problems for research on school choice. Even if students are automatically assigned to a local school and parents are not allowed to choose other public schools, there often exist options for school choice. Students might leave the public school system and attend a private school, provided that there is one within commuting distance. Parents might exercise Tiebout choice, meaning that they move to a residence which is in the assignment area of a desired school. It might also be possible to register the child with someone living in the assignment area of a preferred school, for example the grandparents. Such Tiebout choice is easier to use in areas with a high density of small school districts, where access to several schools can be gained with moves over short distances. In some areas, charter schools are an additional option and some parents might choose home-schooling.

Due to these external choice options, the introduction of an official choice system among public schools is in most cases not the introduction, but an extension of choice options, at least for some parents (Hoxby 2000). The degree of external choice options from these sources varies considerably across school districts. Thus, the actual increase in choice options that results from the introduction of choice among public schools will vary. Moreover, the composition of the student body that is enrolled in public schools is likely to change as a result of school choice. How strong this change will be depends heavily on pre-existing opportunities for choice. There might be, for example, badly performing public schools and therefore many students enrolled in private schools. Private schools usually cost tuition or somehow select their students. Students from households who can afford tuition, and are willing to afford it, tend to be more successful academically. And if schools can select students, by overt or covert means, they tend to cream-skim. As a result, students who are enrolled in private schools are usually more likely to be academically successful. If a high share of students in a school district is enrolled in private schools, this therefore means that the average student who remains in a public school is academically weaker than the average student in the district. If school choice is then introduced, parents of private school students
might reconsider whether access to a private school is worth the tuition. They might not have been willing to send their children to the local public school, but they might apply to other public schools which they think to be of a high enough quality. This kind of reconsideration could result in an influx of students into the public schools, who have on average a higher innate ability than the current student population. If such a change in the student population is not accounted for, estimates of effects of school choice on academic achievement would be biased upwards.

4.2.1.(5) Selection into the treatment

Statistical methods that estimate the effect of a treatment compare the values in the dependent variable of observations that were subject to the treatment to the values for a control group. Selection into the treatment occurs, if the observed entities have an influence over whether they are in the control group or in the treatment group. In the case of school choice, selection into the treatment can occur if choice is optional. By deciding whether to actively choose a school for their child or not, parents choose whether their child will be in the treatment group (=school chooser) or in the control group (non-chooser). This self-selection is a problem for empirical estimates of the effects of choosing schools, if those students who self-select into the treatment differ systematically from those who choose to stay at the assigned school. If choice is optional, those students who apply for a school change are more likely to be academically successful (see 3.1.2.(4)). Even if all students have to hand in a list of schools of choice, as in Charlotte-Mecklenburg County, there are some households who only list the neighbourhood school, thus forgoing their right of choice. These households differ significantly from those who list several schools of choice (Hastings et al. 2005). Thus, students who self-select into the treatment group differ systematically from those who self-select into the control group. A comparison of students who choose schools to those students who stay at the assigned school is therefore likely to be biased. Moreover, some school choice systems are selective. Vouchers are often available only to households who prove that their income is below a given threshold. Students from such households are likely to score below average on tests. On the other hand, those parents who are willing to go through the application procedure care more about academic success of their children than average parents. Students from such households are likely to score above average on tests. The net effect on the average innate ability of students who make use of income-targeted voucher programs is not clear ex ante.
The existence of private schools might also lead to selection into the treatment problems when choice options are extended. Parents who preferred private schools prior to the extension might reconsider if they can now choose among public schools. They might apply to those public schools that they deem suitable, and their children might enter the public school system only if they are accepted at these schools. Students whose parents are able and willing to afford tuition at private schools tend to have a higher innate ability than the average student. Thus, if the introduction of school choice leads to an increase in the observable quality of the public schools, this increase might lead to a migration of students with above average ability from private schools into the public school system. As the low quality and/or prestige of the local public school is one main reason to enter a private school, most students who return into the public school system because of extended choice options would not attend their neighbourhood public school. Thus, these students would be counted as school-choosers. This influx of students would, if it is not properly accounted for, distort most empirical analyses. Analyzes of academic year-to-year gains of schools after the introduction of school choice would be distorted, as the influx of high-ability students raises the average ability of students in the public schools. If this increase in student ability would not be accounted for, it would be mistaken for an increase in educational school productivity. Comparisons between school systems with and without school choice would also be distorted, as extended school choice options attract students with above average ability into the public schools. And comparisons between students who change schools and those who do not change would be distorted, because some high-ability students would only enter the public schools if they were accepted at a popular school. Students who enter the public school system only if they do not have to attend their neighbourhood school would be counted as school choosers. This would increase the average innate ability of school choosers as compared to non-choosers. If not accounted for, this increase in student ability would show up as academic gains that are a result of the school change.

4.2.2. Previous research

Due to the various problems that I have presented in this section, straightforward comparisons across time, across school systems or between school-choosers and non-choosers are likely to yield heavily biased estimates. It is therefore necessary to carefully select a control group, if I want to find unbiased estimates of the effects school choice. I will present the three most promising approaches to identify such a control group, including a few illustrative examples of studies that implemented these approaches.
4.2.2.(1) Approach 1: Lotteries as natural experiments

Several school choice systems admit students to places at oversubscribed schools by lotteries. In these school systems, the students who get access to popular schools should be a random selection of the applicants. And empirical analysis shows that the winners and losers of these lotteries are indeed not distinguishable according to all observable statistics. Those students who got access to popular schools should therefore only differ from those who had to stay at the assigned school by the fact that they won the lottery. Thus, treatment is random and the lottery losers are a perfect control group. This random assignment makes it possible to circumvent the problems that arise from self-selection and systematic selection into the treatment. It is however necessary to be careful if students can apply to several lotteries and I do not have access to the outcomes of all lotteries for all students. In this case, not all lottery-winners will take the places that they have won in the observed lottery, because they have won in another, unobserved lottery, and prefer to take the place at that school. Also, students who lost in one lottery that I can observe might have won in another one, whose outcomes I cannot observe. This student would seem to be a lottery looser but winning the unobserved lottery might have given her access to another popular school.

Which students actually accept a place that they have been assigned by an observable individual lottery is likely to be driven by systematic differences. A comparison of school-choosers to students who did attend their assigned schools or to lottery losers is thus likely to be biased. One way to get rid of selection into the treatment completely in this case, is not to compare choosers to non-choosers, but to compare lottery winners to lottery losers. Assignment to these groups is completely random. On the other hand, such a comparison is likely to dilute the effects of choosing a school, as, as mentioned above, not all winners take the places and some lottery losers might have been admitted at another popular school. Moreover, those schools whose lottery outcomes are unobservable are often no standard public schools. These schools might be private schools, charter schools or special public schools that are allowed to pick their students based on ability. Moreover, some students might only attend the public school system if they get access to a popular public school. Thus, a student who appears to be a lottery looser might have actually attended a selective school that accepted her because of her high innate ability. Or she might have attended a private

19 The probability that a student has successfully applied to a school for which I cannot observe the lottery outcomes depends on the total number of applications that the student has submitted and on the type of schools to which the student applied. As the probability to apply to schools increases with innate ability and SES-background of the family, students who have won in unobserved lotteries are likely to differ systematically from those students who did not win in unobserved lotteries.
school or a charter school that her parents expected to be of a higher quality than the available public schools. If lottery losers can attend private schools or academically selective schools, and if their academic outcomes are included in the dataset, this would lead to a downward bias for the estimation of effects of school choice on academic outcomes.

Cullen et al. (2003 and 2006) used the approach that compares lottery winners to lottery losers. The authors used the outcomes of admission lotteries at high-schools in the CPS. There, students could apply to schools, and places were assigned to applicants by lotteries, that were segregated by ethnicity, gender and grade. Each school held its own lotteries and many schools collected the data that were necessary to conduct the lottery individually. Thus, students could participate in several lotteries independently. This means, as described above, that students might win in more than one lottery or be rejected in one lottery and win in another. In order to circumvent problems associated with selection into the treatment, Cullen et al. therefore compared lottery winners to lottery losers while not using information on whether students had actually taken a place to which they had been admitted. They found that lottery winners attended schools that were better than those of lottery losers in most dimensions on which they had data. But the authors found no evidence that these apparent gains in school quality did translate into systematic academic benefits (Cullen et al. 2006).

This approach, that uses lotteries as a natural experiment to identify control groups, is potentially ideal to identify the effect of being admitted to a popular school on individual students. The assignment into the treatment and the control group is fully random, creating two groups who only differ by the treatment that they receive. A comparison across the groups should therefore show the effects of getting into a popular school, though the effect is likely to be diluted, as explained above. That studies which use this approach often do not show any effects of changing schools on individual students should therefore not be due to the method used. The absence of any significant effect when using this method can then be either because there are no such effects, because the effects are weak and diluted below significance level, or due to the fact that most parents are not able to identify the more productive schools, as described above. But even if there exists an effect in reality, and parents chose the highly productive schools, this approach would not be capable of identifying overall effects of choice. An overall increase in educational productivity due to competitive pressure from

---

20 As mentioned above, there might be problems if it is not possible to observe the outcomes of all lotteries and the academic outcomes of students who attended schools whose admission procedure cannot be observed is included in the estimation.
school choice would be obscured, as both the treatment and the control group would benefit from such an increase.

4.2.2.(2) Approach 2: Variation in external choice

Another strand of literature uses an approach that is more capable of identifying overall- and long-term, general equilibrium effects of school choice than the comparison of lottery winners to lottery losers. This approach uses variations in the extent of choice options that parents have, before they are allowed to choose among public schools. As described above, such choice options could be offered by private schools, in combination with vouchers or not, by charter schools or via Tiebout choice. If more choice opportunities from any of these sources are available to parents, public schools risk loosing more students. Even if there are no explicit sanctions for loosing students in place for public schools, reputation effects and the likely reduction of staff positions, or even the closure of empty schools, are likely to result if students desert a school. The resulting pressure should induce public schools to work harder if parents have more outside-options. Several studies have taken advantage of variations in the availability of external choice options to identify overall effects of school choice. As the extent of choice from these sources does not change quickly, this approach is capable of identifying long-term effects of school choice.

Reverse causality is a problem for this approach, however. Both charter schools and private schools are more likely to be found in areas with bad public schools that hold dissatisfied parents (Hoxby 2003b). Simply comparing public schools in areas with many charter schools to public schools in areas with few such schools, for example, would therefore yield an estimate of the effects of competition that is biased downwards.

It is therefore necessary to find exogenous determinants for the number of schools that compete with the public school system. One attempt was made for private schools in the US. There, a large share of private schools are confessional schools, mostly Catholic ones. Many of these schools admit students who do not share the faith of the school and they often charge tuition fees well below actual costs. Therefore, they are a viable competition for a large part of the clientele of nearby public schools and their presence should result in pressure on public schools. But the density of Catholic schools might be affected by the demand for them. And the propensity to attend a Catholic school might be driven by characteristics that also affect educational attainment. Catholics are for example much more likely to attend a school of the same denomination. And Catholics in the US are more likely to have a south-American, south-European, Irish or Polish background than the average student. Apart from possible
cultural differences regarding education across ethnic background, the last generation to be born outside the US is certainly correlated with ethnic background and likely to affect academic achievement. Thus, the density of Catholic schools might be a reaction to the share of Catholic students in the area, and these might differ systematically from students of other denominations.

Therefore, a comparison between areas with a currently higher share of Catholics, or students enrolled in Catholic schools, to areas with lower shares or to students enrolled in other schools, might be distorted by systematic differences in the local population that also affect outcomes. Most Catholic schools were however founded decades ago, and tend to survive for a long time. Therefore, Catholic schools are more common in areas which happened to have a large Catholic population in the past, which provided the means to found these schools. Thus, the share of Catholics in the population at an earlier point in time was proposed as an instrumental variable for competition to public schools from Catholic ones (Hoxby 2002a).

Hoxby used the share of major religious denominations in 1950 and controlled for current religious composition. Using this approach the author found a significant and substantial effect of competition from choice options at private schools on the productivity of public schools.

Charter schools in the US provide another possible approach. The founding of charter schools is heavily regulated. The number of charter schools being founded is thus not only a reaction to demand for this type of school but also a reaction to rules about founding charter schools. And these rules vary considerably across constituencies. Thus, a possible instrument for the existence of competition from charter schools could be a measure for the difficulty of founding such schools. However, these rules are certainly not created randomly. Teacher Unions usually oppose school choice. In constituencies where teachers have more reason to fear competition, this opposition should be more determined and thus the permitted competition less intense. As competition is more dangerous to teachers if parents are disappointed with public schools, this raises the possibility of reverse causality problems.

There are two strategies that have been used to deal with the reverse causality problem when analyzing the effect of competition from charter schools on public schools. One strategy is to search for variations in the ease of founding charter schools that are not driven by the demand for them. Eric Bettinger found such a variation in Michigan (Bettinger 2005). All charter schools need to be approved by some authority. In Michigan, state universities were allowed to approve charter schools. The boards of a majority of these universities were appointed by Governor Engler, who was an avid supporter of charter schools and used his influence to
convince these universities to approve charter schools. Founding a charter school requires a lot of contact with, and supervision by, the approving authority. It should thus be easier to found a charter school if a state university is nearby and the numerous meetings that precede a charter approval do not involve extensive travelling. Bettinger tested whether the distance of a public school to one of the 15 state universities could be used as an instrument for the likelihood that a charter school was established nearby. He found that this distance was a viable instrument. Using this distance to instrument for charter school competition, Bettinger found little or no effects of a nearby charter school on test scores (Bettinger 2005).

The other strategy takes a different approach. It tries to circumvent the problem that charter schools are more likely to be founded under certain circumstances by analyzing changes within schools. Hoxby (2003b) analyzed changes in the achievement within existing public schools that were triggered by newly arising competition from charter schools in Arizona and Michigan. To do this, she divided public schools into two groups. The first group consists of schools that were confronted with new competition from charter schools. These schools were identified by the fact that the share of students in their vicinity who attended charter schools crossed a specific threshold. The second group consists of public schools in areas where charter school enrolment did not cross this threshold. Hoxby then used a difference in differences approach that compares changes in the performance of the two groups of public schools over time. While doing so, Hoxby tried to account for systematic differences in the unobserved variables that might affect the performance of schools and the competition that they face from charter schools. The author did so by using school fixed effects that account for time-constant unobserved characteristics. Additionally, she analyzed changes in trends of student achievement to account for trends that are independent of the new competition from charter schools. Hoxby found significant improvements in fourth grade math and reading of public school students as a reaction to charter school competition. She also identified an increase of the annual growth rate of achievement as a reaction to competition from charter schools (Hoxby 2003b).

As illustrated in the two examples above, findings using this approach differ considerably. This might be due to the heterogeneous nature of charter schools. There are for-profit schools founded by companies that specialize in this industry. Some charters are founded simply to provide good schooling in an area with failing public schools. And there are charter schools founded by groups of parents or teachers who want to follow a certain educational philosophy. This philosophy might be efficient in preparing students for tests that assess achievement. But it might also focus on goals that do not directly benefit achievement as
measured by common aptitude tests. Examples for other goals would be to encourage students to think independently, to focus on arts or to improve their ability to resolve conflicts. It is likely that the mix of motivations of charter school founders varies, both across school systems and across time. This could lead to differing results if the used data are taken from different regions or different points in time. Another issue that is related to timing is the age of charter schools. Districts in which more than a small fraction of students is enrolled in charter schools are a quite recent phenomenon. And most of these schools are only recently set up and rather small. This means that they might have a selective intake which favours children of the founding members. These schools might also have starting problems. This problem might, for example, be financial, because the new staff has to learn working together or because the administration has to learn how to run a school. Additionally, the enrolment is in most areas too small to seriously challenge public schools. The age of charter schools might explain the different findings of Hoxby and Bettinger. While Hoxby had access to data up to 2002, Bettinger conducted his analysis in 1999, using mainly data from years in which most charter schools were recently set up and did not contain a big enough share of students to put public schools under serious pressure.

A third approach, also developed by Caroline Hoxby (Hoxby 2000), uses variation in the size of school districts. The smaller the districts, the higher their density and thus the more variation there is regarding schools within a given area. With more districts in a given area, it is easier to move into the assignment area of schools that are run by another school authority. The size and number of school districts might however be due to systematic differences that are correlated with monetary school productivity\textsuperscript{21}. In the US, voters resist the consolidation of school districts when one or more of the districts that are to be merged show a bad monetary productivity. If, on the other hand, the districts share a high monetary productivity, voters tend to support consolidation in order to reap benefits of scale. Also, in districts with a low school quality, neighbourhoods with a relatively better school quality try to break away

\textsuperscript{21} In the studies that are presented in this subchapter, Hoxby uses monetary productivity, which is measured as outcome score per dollar that is spent on student financing. This productivity is different from educational productivity, which I use throughout this dissertation and which is measured as outcome score for a given set of student and school characteristics. Educational productivity isolates the contribution of the school to the academic attainment of students. It increases only if the school gets better at educating students. Monetary productivity can also increase if teacher salaries are reduced or if the student intake of one year has a higher average innate ability. Hoxby usually uses difference-in-differences estimations that compare schools with similar or identical financing schemes and takes care to compare schools with very similar student characteristics. In this case, where financing is very similar and systematic differences in the student intake are controlled for, monetary and educational productivity are almost identical.
and found their own district, while such attempts at secession are rare in districts with a high school quality (Hoxby 2002a). These observations imply, that the density of school districts might be driven by the monetary productivity of districts and therefore be endogenous. Thus, it is necessary to find a measure for the density of school districts that is not affected by the demand for Tiebout choice. Hoxby found such a measure by exploiting the fact that the boundaries of school districts in the US were initially set in the eighteenth and nineteenth century and have not changed a lot since then. At the time when school districts were defined, the most important concern in setting the boundaries was the time that students needed to travel to reach the next school. This was mainly determined by travelling distance which in turn was strongly determined by natural barriers that were hard to cross. The most important type of such natural barriers are streams, which used to be formidable obstacles to movement, at least for parts of the year, at the time when boundaries of administrative areas were determined. These streams have meanwhile often been rerouted, dammed or bridged, or have even disappeared underground for long stretches, so that they are much less of an impediment to travelling today. But the boundaries that were defined alongside these streams have often survived unchanged. Hoxby therefore tested the number of streams in an area and found them to be a viable instrument for the extent of Tiebout choice in an area (Hoxby 2000)\textsuperscript{22}. Using this instrument, the author found significant and sizeable effects on monetary school productivity. An increase from 0 to 1 on the index of Tiebout choice, which represents the difference between a large metropolitan area with only one district, and another with several dozen districts, raises test scores by one quarter to one half of a standard deviation. Average overall effects are moderate however, as the standard deviation for the choice index across metropolitan areas is 0.27. Although moderate, these effects are still much stronger than those identified with a simple OLS regression, which ignores possible demand-side effects on the Tiebout choice index. Such an OLS approach shows no, or very small, effects of Tiebout choice on achievement (Hoxby 2000). This difference in findings highlights the importance of

\textsuperscript{22} The approach and findings of Hoxby (2000) were criticised by Rothstein (2005). Rothstein criticized, among other things and in an almost insulting tone, the construction of the measure for streams used by Hoxby and the fact that private schools were excluded. Rothstein also used a different dataset, claiming that Hoxby had refused to make the data she had used available. In an unusual reply to the comment of Rothstein, Hoxby (2005) stated in the abstract that “every claim of Rothstein was wrong”. Among other things she stated, that her original raw data were publicly available as well as the code she used on this raw data and that she did not provide temporary intermediary datasets Rothstein had asked for as these had not -and should not- be stored. Hoxby also defended her construction of the stream variable. She also stated that her results are robust to the inclusion of private school students. According to Hoxby, Rothstein did not only add private school students but additionally used different data, substituting error-free data with error-prone data and that this change in data was the reason for the different results that Rothstein had come up with. Being aware of both the original paper, the comment of Rothstein and the reply of Hoxby to this comment, I come to the conclusion that I can trust the general findings of Hoxby (2000).
using a carefully developed identification strategy when analyzing the effects of school choice.

The insights gained with this approach indicate, that increased competitive pressure, which results from choice options outside the public school system, has an effect on school behaviour. Public schools react to increased competition from outside-options by increasing monetary productivity.

But such cross-sectional approaches, that analyze the average effect of the degree of school choice on many school systems, only identify effects that are averaged across many school systems. Moreover, the strength of the identified effects is often diluted by the necessity to use instrumental variables. This approach is therefore not capable of predicting the effects of an introduction of expansion of choice on a specific school system. Moreover, such cross-sectional approaches are not capable of giving insights into the effects that specific institutional characteristics and conditions have on school choice and its outcomes.

4.2.2.(3) Approach 3: The introduction of school choice as a natural experiment

The introduction of school choice is a policy decision. Such a policy is often the outcome of pressure for choice. And it is often implemented in a way that makes it hard to use this change in policy in order to analyze the effects of choice. But if school choice is introduced in a rather small area, and if there is a reasonably similar school system with a rather similar student population close by, such an introduction of school choice can be exploited as a quasi-experiment. The optimal case is the introduction of school choice in a randomly chosen part of one school authority. In this case, there likely are student populations close by that do not differ from those populations who are now allowed to choose, and who are subject to identical conditions except the possibility to choose schools. The similarity of conditions is important, as differences in conditions, like a different pool to draw teachers from, different funding of schools or a student population that differs systematically can distort estimates. Under optimal conditions, the effects of school choice can be estimated quite precisely by comparing the treatment group to the control group.

The best example for such an introduction of school choice that was documented so far could be observed in Tel Aviv in 1994. There, an existing system of busing students to assigned schools was replaced by a school choice system. Students were each allowed to choose among 5 schools, two within their district and three in another district. Importantly, this school choice was first introduced only in one of several school districts of Tel Aviv, which
are all administered by the same school authority, and therefore share most institutions and regulations. Moreover, this introduction of school choice was the only change introduced at this time\textsuperscript{23}. The result was a rare situation, in which the introduction of school choice is the only major change at a time. Moreover, there was a student population who was subject to conditions that remained identical to the conditions that had been faced by the treatment group before treatment began. The district in which school choice was introduced was a rather poor one, with a high share of immigrants. To obtain a control group with characteristics similar to those of the treated students, Victor Lavy only compared students who were living within a band of 200 to 250 yards on both sides of the district border. This yielded a group of students who did not differ significantly from the treatment group on any characteristic for which data was available. Using a difference-in-differences approach, the author compared changes in outcomes of the treatment group over the course of several years to changes in outcomes of the control group (Lavy 2006).

The particular combination of the set-up of this reform, the chosen control group and the econometric methods used by Lavy, are capable of dealing with all major difficulties of conducting research on school choice that were described above. Self-selection, systematic selection into the treatment and reverse causality pose no difficulties as choice was only available to students residing in one district. The control group was identical to the treated group and did only not participate because it lived a few dozen yards away, on the other side of an arbitrary administrative boundary. External choice options were the same as, again, the identified treated and control group lived closely together. There were no other reforms conducted at the same time. And, as only one school system is analyzed, variations in institutional settings are no issue. This approach is thus perfectly capable of identifying the effects of the introduction of school choice in Tel Aviv.

Unfortunately, this approach is capable of identifying the effects of \emph{this particular kind} of school choice in this very school system only. As discussed at length in chapter 3.3., various institutional settings can have strong effects on the outcomes of school choice. The characteristics of the student population also have an effect. Unless there exists another school system, in which a similar system of busing is replaced by a limited choice between schools within and without the home district etc., insights of the study of Lavy cannot be applied easily to predict the resulting effects in another school system. Moreover, it is not possible

\textsuperscript{23} There were of course some flanking measures that have to be decided upon when school choice is introduced. For example, what is to happen to teachers at schools that loose students? Most important among the decisions taken in Tel Aviv is the decision to close down schools that loose too many students. But there were no other reforms introduced at this time that were not necessary in order to implement school choice.
with this approach to identify the effects of individual institutional settings, as no variation of settings can be observed with only one school choice system.

### 4.2.2.(4) Summary on empirical approaches

Regarding the strengths and weaknesses of the three approaches that I have presented above, it is clear that empirical research on the effects of school choice will always generate limited results. Each of these approaches, no matter how well they are done, can only identify insights that are limited in one of two possible ways. An approach can either identify detailed effects of one particular kind of school choice on one particular school system. Alternatively, an empirical approach can identify average effects of a given type of school choice across many different school systems. Both types of approaches help to get an idea about the potential of school choice to change school outcomes. But neither is capable of identifying effects of individual institutional settings on the outcomes of school choice. There is a wide range of settings in which school systems and school choice regulations differ. These settings can heavily affect the outcomes of school choice. Therefore, understanding the influence of such individual settings is vital to predict effects of school choice - or of changes in existing school choice regulation - on the outcomes in given school systems.

Studies like the one on Tel Aviv have the potential to get insights into the effects of individual settings eventually. If a large enough number of studies of this type and quality would be available, and if characteristics and settings in all important areas of the analyzed school choice systems were known, a meta-study would be a possibility to get insights into the impact of individual settings. Unfortunately, authorities who decide on the introduction of school choice rarely present researchers with such good conditions as Lavy found in Tel Aviv. And in those studies that already exist, institutional settings are often not sufficiently documented for a proper analysis. Moreover, given the wide variability of settings in school choice systems, a great number of studies would be necessary to obtain significant estimates. The possibility of strong cross-effects of individual institutional characteristics would necessitate interaction variables for settings in these characteristics, which would further increase the necessary number of observations.

Another drawback of empirical approaches is, that these approaches can only investigate variations and combinations of institutional settings that have already been implemented in a school choice system. Theoretical approaches, which will be presented in the next subchapter, could however be capable of predicting the effects of institutional settings that have not been documented before.
4.3. Theoretical approaches

In this section, I will analyze characteristics of school choice that lead to problems for theoretical approaches to research in this field. I will start by presenting these problems (4.3.1.) and then present the computational general equilibrium (CGE) approach, which avoids most of these problems. I will also present a few studies which were among the first to apply the CGE approach to school choice (4.3.2.). While being an innovative and important step forward, the CGE approach and the studies that have applied this approach so far have a few weaknesses, which I will describe in the last section of this chapter (4.3.2.).

4.3.1. Problems for theoretical approaches

4.3.1.(1) Interdependence of decisions

The main problem that theoretical approaches encounter when applied to school choice is the interdependence of decisions and actions across actors. This interdependence was already mentioned in 4.1.(2). School choices by parents affect, via a changed student body, the outcomes of decisions that are taken by other parents. As a consequence, these parents might want to reconsider their previous decisions. School choices by parents also affect decisions of schools, as they change the number of enrolled students and aggregate student characteristics. And decisions by schools change test results and thus affect choices of parents and the competitive threat that individual schools pose to other schools.

These interdependencies often act as amplifiers for other effects. If, for example, school A, which has an academically weak student population, raises educational productivity, this school then achieves better academic outcomes. The improved outcomes attract additional applications, so that school A grows. As high-ability students are more likely to apply to any school, the new students of school A will have a higher average innate ability than the former student population. This influx of high-ability students will thus further improve average academic outcomes of school A. Thus, the effect of a one-time increase in educational productivity on average academic outcomes is amplified by the reactions of parents.

Note here, that the reactions of many parents that followed the initial increase in educational productivity were no direct reactions to the change in educational productivity of school A. Many parents reacted to a later change in the academic outcomes of school A, which was caused by previous reactions of other parents to the initial increase in educational productivity. Moreover, the total improvement of school A concerning academic outcomes might induce an increase in effort by school B, that competes for the same students. This
reaction, again, is either triggered or amplified not by the initial cause (the original one-time increase in educational productivity of school A), but by reactions of other actors (parents) to the initial cause.

Depending on the measure that parents use to assess the academic quality of schools, the interdependence of decisions can even create strong dynamics without a real cause. One example is, when a school in an area with a very weak student intake is closed, for example because of underutilization or consistent underperformance. The students of this school need to be accommodated elsewhere. As most parents dislike long commuting distances, many displaced students will apply to schools that are close to their home location and thus close to the old school. Students who attend schools that are closed tend to be academically weak on average. Those who then attend nearby schools, instead of applying for better schools that are farther away, tend to be even weaker academically. A strong influx of students from a closed school is thus likely to decrease average academic outcomes and worsen the average characteristics of the student intake. Even if these changes do not affect their own children (assume that all replaced students are concentrated in a new grade of the school), parents who strongly care about academic achievement are likely to desert schools that have accommodated many displaced students. As these leaving students tend to do better in school, results and average statistics at the asylum-schools drop even further. Thus, the school that receives replaced students seems to undergo two decreases in schools quality, although the educational quality of the school is unaltered.

Many theoretical models use representative agents. It is however not possible to catch the effects of the presented interdependencies with representative agents. Parental school choices are for example strongly driven by average school characteristics which are in turn affected by the decisions of schools and other parents. And considerable shares of the effects that eventually result from an original triggering event are reactions to previous reactions of other actors. This has also implications for the timing of effects. If a large share of the eventual effect of a change is due to reactions to previous reactions of other actors, the full effect might take a long time to materialize. Ignoring the interdependencies would miss a large part of what is actually going on in a school system and would thus heavily distort any findings.

4.3.1.(2) Heterogeneous behaviour and characteristics

I have already presented the problems that arise when the heterogeneity of actors is ignored in 4.1.(3). Most theoretical approaches use representative agents to take the place of a type of actor, like the students. It is also possible to use two or more different representative agents,
who each represent a group of actors who differs in an important characteristic or preference from other actors of the same type. In school choice however, students and schools differ in several dimensions that strongly affect both their actions and the effect of their decisions. High-ability students are, for example, more likely to change schools and are in most settings more beneficial to the receiving school than low-ability students. Independent of innate ability, students from a low SES-background are less well informed and less likely to change schools than better informed students from a high-SES background. The household location affects the distance to all schools and thus the propensity to apply there. And schools are affected by the characteristics of the student population that lives close by.

Using more than two types of representative agents for one type of actors (here: schools or students) tends to result in messy models that are hardly tractable and have difficulties to produce clear-cut results. Therefore, such models are not often used. Students are heterogeneous in several dimensions that have a strong effect on their choice behaviour and the outcomes of their decisions. To represent these differences appropriately would necessitate considerably more than two types of agents. Schools are also heterogeneous in dimensions that affect outcomes. Examples are the quality measure and the location of schools. Both characteristics affect the attractiveness of schools to students. A theoretical model in which both students and schools vary across several dimensions is likely impossible to solve only with pen and paper.

### 4.3.2. Computable general equilibrium approaches

#### 4.3.2.(1) Introduction to CGE approaches

One way to account for heterogeneous agents and the interdependency of decisions is to use computational approaches. In these approaches, the actors are generated or taken form real world data, and can have differing characteristics, information and preferences. In the case of analyzing school choice, these actors are individual students and schools. To give an example for the characteristics of an actor in such a computational model: a household is located at a given address, so that travelling distances to surrounding schools can be determined. This household has a moderately high SES and holds a Caucasian boy of average innate ability who speaks fluent English.

In a CGE model, decisions are determined for all actors individually, taking into account their specific characteristics, information and preferences. This means of course, that one maximization problem has to be solved for each actor. Even when using only a few hundred students and a dozen schools, the resulting workload used to be impossible to handle. Even a
decade ago, the computations would have necessitated expensive computing time at a mainframe. But specialised software and technological advances that reduced the price of computing power have lead to a situation in which CGE models can be handled by personal computers.

While the heterogeneity of actors is easily taken care of by using differing characteristics and by solving one maximization problem per actor, dealing with the interdependence of decisions necessitates additional thought and effort. The outcomes of decisions of individual actors are affected by the decisions that are taken by other actors. Consequently, many actors might want to reconsider their decisions once they have learned about the outcomes of the decisions taken by other actors. And if actors can make predictions about the decisions of other actors, they are likely to consider those predictions when taking their own decisions. The CGE approach tackles the difficulty of interdependent decisions by using repetitions of the decision process. In a first step, all actors take their decisions simultaneously, assuming ceteris paribus conditions. Then, outcomes are calculated. These outcomes change the information on which actors had based their decisions. In the next step, the information of all actors is updated to represent the outcomes of the decisions in the first round. Then, the maximization problems for all actors are solved again, based on the new information. Once this second maximization is finalized, results are calculated, information is updated and actors decide again.

An illustrative example of the mechanism of the steps in this maximization procedure is the following: a school is populated by local students with a high average innate ability and SES. This high average SES attracts many applications. From these applications, a lottery admits a random selection of students with an average SES that is below. As a result, the average post-choice SES at the school decreases. After noticing this new level of average SES, which is a result of simultaneous decisions by other actors, many parents who would have chosen this school based on the old average SES might want to reconsider. They now decide to apply to other schools instead. The lottery assigns students again to schools, parents observe aggregate characteristics of the student populations and some reconsider their choice decision again.

This process of decision-making, computation of results, updating of information and another round of decision-making is repeated until, after several rounds of incremental convergence, a general equilibrium is reached. This computational general equilibrium is characterized by a situation in which no, or only a small share of the actors would want to reconsider their decisions given the changed information that resulted from altered decisions of other actors. This situation is a general equilibrium, as the vast majority of actors, who are students and
schools in the case of school choice, do not want to change their decision, given the decisions of the other actors. In the extreme case, in which the share of actors who is allowed to change decisions is set to zero, such a general equilibrium is identical to a Nash-Equilibrium in pure strategies: given the decisions of all other actors, no actor does want to change his decision as he cannot improve his pay-off by doing so.

It is important to note here, that the general equilibrium is usually assumed to be reached instantaneously. Repetitions of the decision process do not represent turns, like years in the case of school choice. These repetitions are only intermediate steps of a computational algorithm that is used to identify a set of decisions with characteristics that constitute a stable state. A CGE could in reality be reached instantaneously, if all actors were capable of predicting the decisions of all other actors perfectly. For an instantaneous CGE, these predictions would have to consider the effects of own decisions on the decisions of other actors, i.e. an actor would have to know that the decisions of other actors depend on his own decision and that the other actors can perfectly predict his own decision. The assumption of universal perfect predictions is unrealistic, however, so that a CGE is very unlikely to be reached instantaneously in a real school choice system.

This use of computational power to simulate the behaviour of many actors is an important and innovative step forward in the research on school choice. It makes it possible to deal with the problems that arise from the heterogeneity of actors and the interdependence of decisions. Moreover, the use of a general equilibrium approach deals with problems that are due to the fact that changes in the first years after the introduction of school choice might be misleading (see 4.2.1.(2)): once a general equilibrium is reached, all effects of school choice have materialized fully and all students have had a chance to make use of the new choice options, not only those students who tend to move first.

Pioneers in the use of CGE approaches to analyzing school choice were Epple/Romano (1998), Nechyba (1999) and Fuchs (2007). I will proceed by presenting their work.

4.3.2.(2) Example one: Epple/Romano (1998)

The first researchers to apply a computational model with many actors to school choice were Dennis Epple and Richard Romano in their 1998 article in the American Economic Review. They start by setting up a model that contains many students and schools. Students differ by family income and innate ability. Both these characteristics of students are assumed to be observable and a positive correlation is assumed between family income and innate ability. Schools are either tax-financed, tuition-free public schools or private schools that charge
tuition. Public schools are identical and have to accept all students, while private schools strive for profit and can enter and leave the educational market. Private schools charge tuition dependent on ability of the student and income of the household. By using this price discrimination, they can steer the ability-mix of their student intake. This mix is important, as the educational attainment of a student is a function not only of the innate ability of an individual student, but also of the average ability of her peers.

Epple and Romano start analyzing this model by using a traditional comparative-static method. Then, they develop a computational version of the model in order to illustrate the results of the first method, to analyze the effect of vouchers and to investigate issues on which the comparative-static method yielded ambiguous results. This use of computational methods is necessary as the model is too complex to be tractable with non-computational methods and to generate clear-cut results on many issues.

The authors use simulations based on their model in order to analyze the effect of vouchers on sorting. They find a CGE which contains a strict hierarchy of private schools which is ordered by quality and, by assumption, a homogeneous public school sector. Students are sorted into these schools according to their characteristics, innate ability and family income. School quality is modelled by the authors to be affected by economies of scale. This makes small schools unprofitable and thus prevents an equilibrium in which a vast number of schools cater to infinitely refined peer groups. In this equilibrium, private schools offer tuition discounts to low-income, high-ability students and let low-ability, high-income students pay for the increase in school quality that results from having high-ability peers. Regarding vouchers, the authors find that higher voucher values increase the relative size of the private sector, increase sorting and increase the gap of school quality between high-ability students and low-ability students.

The use of empirical findings and real world data is rather limited in the calibration of this model. The authors use the US Census to obtain the distribution of family income for the students. To obtain the ability distribution, the authors start by assuming that income is determined by educational achievement. Individual educational achievement is in turn assumed to be determined by own ability and peer ability, following a Cobb-Douglas function. Epple and Romano assume, loosely based on empirical evidence of previous research, that the elasticity of achievement with respect to peer ability is 20% of the elasticity of achievement with respect to own ability. Thus, if the elasticity of educational achievement with respect to own ability is known, the achievement production function can be determined. Due to the nature of the assumed utility function, this elasticity is identical to the elasticity of
utility with respect to school quality. The preferences of households regarding school quality are determined in the following way: given a lack of empirical evidence on the demand for quality, the authors assume unitary price and income elasticities. This allows them to set the utility elasticity of school quality equal to the share of disposable personal income that is spent on education in the US, which is about 0.06. The elasticity of achievement with respect to own ability has, as a result of previous assumptions, to be five \((1/0.2 = 5)\) times the elasticity with respect to peer ability. It is thus set at 0.3. The authors then calibrate the ability distribution so that the educational achievement function that was derived above generates, in the steady-state equilibrium, an income distribution that matches the distribution that was documented by the US census.

Differences in the educational productivity of schools are assumed to be non-existent for simplicity. This means that all schools will generate the same average educational achievement, if given the same student intake.

4.3.2.(3) Example two: Nechyba (1999)

The second study to apply a computational model with many actors to school choice was conducted by Thomas Nechyba and was published in 1999 in the Journal of Public Economic Theory. The author defines a model with multiple school districts. Each of these districts is financed by property taxes which are set locally through a majority vote. There are public and private schools. Students have no choice among public schools across district boundaries. Private schools are clubs of parents that can exclude students while public schools have to accept all students. Educational achievement in this model is computed with a Cobb-Douglas function which is determined by average peer ability and spending per student. Households differ by their endowment, which consists of a house, an amount of a private good and their innate ability. The household utility function increases with the quality of the community in which the household lives, consumption of the private good, the quality of the house that is owned by the household and the quality of the school that the child attends. This school quality is determined by per pupil spending and average peer quality, which is the average innate ability of the attending students. It is important to note that Nechyba does not take a stand on the shape of the actual educational production function. He sees school quality simply as a measure that parents have at their disposal and on which they base their school choice decision. Thus, “school quality” in this model is the quality \textit{as perceived by the parents}, and does not carry any information about educational school productivity.
As in Epple/Romano (1998), the theoretical model is too complex to yield many clear-cut analytic results. Therefore, Nechyba develops a computational counterpart of the theoretical model. The calibration of this simulation model does not rely strongly on real world data: preferences of households for the desirability of living in individual neighbourhoods are calibrated to real world data (from New Jersey). Household income can take on one of five values, house endowments can be of nine different types. In combination, these characteristics create 45 endowment types, each of which is assumed to hold an equal share of households. Innate ability is assumed to be perfectly correlated with income. The relative impact of per pupil spending and average peer quality on school quality is set to 0.5 in most simulations, based on an ad-hoc assumption.

Nechyba then uses computational simulations, which are based on the presented model, to investigate the impact of vouchers in the context of different regimes of education financing that had been in place before vouchers were introduced. He finds that vouchers would benefit public schools in poor communities and hurt public schools in wealthy communities.

4.3.2.(4) Example three: Fuchs (2007)

Following Epple/Romano (1998) and Nechyba (1999), Fuchs applied a computational model to school choice in his dissertation. There are two major differences to the two approaches described above. The first difference is, that the model of Fuchs is directly aimed at a computational approach, unlike Epple/Romano and Nechyba, who start with a theoretical model and use the CGE-approach only in a second step to get insights that the theoretical model cannot provide. The second difference is the inclusion of a spatial dimension. Fuchs assumes an imperfectly mobile population with fixed locations for student residence and schools and mobility costs for students. The effects of this assumption are a preference for schools that are closer to the home of an individual student and that the characteristics of the local student population have an impact on school characteristics. The inclusion of the spatial dimension is meant to be an adaptation to the conditions in Europe, where Tiebout-choice is used less frequently. Concerning model mechanics, the spatial dimension puts an end to the footloose characteristic of students who will go to the school that best fits their preferences, no matter where it is located.

The model of Fuchs is based on the model of Epple/Romano (1998), enriched by the spatial component and a specification from Nechyba (1999) that includes an effect of the socio-economic background of peers on educational achievement. Households are endowed with a family income, a level of student ability and a socio-economic status. The utility-function of
the household depends on disposable income, student achievement and mobility costs from attending a distant school. Student achievement is a function of student ability and school quality. School quality is a function that increases with average peer ability and the average socio-economic status of peers and decreases with school size. This means that increases in the quality of individual schools can only be an effect of better peers or a smaller school size. Increases in average school quality for the entire school system can in this model only be induced by a different distribution of students to schools that minimizes losses from school size and maximizes gains from the combination of peer characteristics. Schools cannot affect their productivity directly in the model of Fuchs and the school quality is perfectly observable to parents.

The specification of the model mostly follows Epple/Romano (1998) and is only rudimentarily based on empirical foundations. Concerning the spatial dimension, the distribution of households is generated randomly, with a denser population in the centre of an assumed city.

Fuchs uses this model to analyse different school choice mechanisms and school financing regimes that involve free public schools and costly private schools. The tuition of the private schools is non-discriminatory or discriminatory, depending on the setting. Fuchs uses a CGE-approach to analyze the impact of the use of two types of vouchers, flat rate vouchers and targeted vouchers, in several scenarios. He finds that school choice increases achievement by reducing inefficiencies that result from an unfavourable student distribution across schools. He also finds that the highest increase results from a schooling market that only consists of non-discriminatory private schools and that this scenario also produces the highest level of educational equality. His final important result is that discriminatory tuition strongly reduces educational equality.

The CGE-simulations of Fuchs are done using a model programmed in STATA (Stata 10). This CGE model was the starting point for the programming of the first version of my own micro-simulation model that I will describe in chapter 6. I greatly expanded the model and changed the existing features. Among other things I adapted the model to be based on real world data, vastly increased the number of school and student characteristics and estimated the reaction functions of schools and students. Also, school productivity can be affected by schools, based on an estimated education production function, instead of being affected only by student characteristics. As a result of modifying and expanding the model, I would be surprised if even one line of code of this original simulation of Fuchs survived unchanged in the model that I used to generate the results in Chapter 9. But the simulation of Fuchs was
invaluable for my work as a starting point and provided some crucial ideas on how STATA could be used to program a micro-simulation model of school choice.

4.3.3. Weaknesses of previous CGE approaches

While being an important step forward, previous CGE approaches to analyzing school choice share a few weaknesses. These weaknesses do not seriously affect the analyses that the authors conduct with their respective CGE models. But they would seriously hamper any attempts to use such CGE models in order to achieve insights into the effects of most institutional features of school choice systems. These weaknesses are partially due to the nature of the CGE approach itself and partially due to a lack of data and calibration in those models which were used by the authors of previous studies.

4.3.3.(1) Weaknesses of the CGE approach

The major weakness of the CGE approach is a result of the general equilibrium nature of all findings. This equilibrium is, as described above, the outcome of repeated iterations in which individual actors change their decisions incrementally, until a situation is reached in which no actor wants to deviate. To reach such an equilibrium, it is often necessary to repeat the maximization procedure several dozen, or even several hundred times. During this procedure, an individual actor might change his decisions significantly. And he might change these decisions multiple times. The updating of information, and thus an opportunity to change decisions in face of this new information, does not occur frequently in real school choice systems. Usually, new information on school characteristics, and another opportunity to change schools, is presented yearly. This means, that such an equilibrium might only be reached after decades, or even centuries. Moreover, the actors in real school choice systems are constantly replaced. Households leave the decision process when their children graduate or drop out, and new households enter the process when their children enter school. This constant replacement of student populations, and its effect on aggregate student characteristics of schools, ensures that a stable equilibrium is never reached. And most students would be unwilling to change schools each year. Thus, in order to avoid another school change, many students might stay at schools that would not have been their first choice, given the newest information on school characteristics.

To assume that actors can predict the decision of other actors could alleviate the problem of an unrealistically high number of iterations. By making predictions that come close to the actual decisions of other actors, individual actors can reduce the probability of changing their
own decision, once they know the decisions of all other actors. Such predictions would reduce the number of repetitions that are necessary to reach a CGE. But empirical observations show, that parental school choice decisions are mainly based on quality measures, and other characteristics of schools, which were documented for previous cohorts of students (see 3.1). Moreover, even rough predictions of parents and schools about the decision of all other parents and schools would require considerable time and effort. And even given this effort, predictions are likely to be imprecise. Therefore, it is likely that predictions, if made at all, would be rudimentary and would mainly be based on the latest known characteristics of schools. Examples for such predictions could be: (A) In the last years, there were many minority students at school A and it had test outcomes that were below the average. I therefore assume that there will be many minority students next year as well, and that test outcomes will be below average again. (B) School B has shown considerable increases in test outcomes last year. I assume that a lot more parents will apply to this school, and many of those applicants will be from our own school, as it is close to school B.

Summing up, predictions could reduce the number of repetitions that are sufficient to reach a CGE in the school choice case. But predictions are unlikely to reduce the number of repetitions to a level that could be reached within a few years, especially given the constant replacement of the student population and an unwillingness to make repeated school changes.

Another weakness of the use of a general equilibrium, that might or might not materialize at some distant point in time, is that the patience of school authorities is usually limited. When considering the implementation of a school choice system, or a change in the regulations of an existing one, it is more important to know what would happen in the next few years. A policy that fails to deliver results within a few years is likely to be altered or even revoked. Therefore, the eventual outcome of such a policy in a few decades is not of immediate interest to school authorities, especially as this state is unlikely ever to be reached. The transition path to any kind of fairly stable state is, on the other hand, of immediate interest to school authorities.

A new policy that would result in strong improvements in student attainment for all students in 20 years, but would mean strong losses for some non-negligible share of the student population meanwhile, might not be left in place long enough to bear fruits. An example for this would be the introduction of rather free school choice with rather complicated application procedures and drastic consequences for deserted schools. This combination of free choice and strong consequences for schools which loose many students could eventually lead to a
higher average effort of teachers and thus to an overall increase in educational school productivity. But school-choosers tend to be academically stronger than non-choosers. And due to the complicated application procedure, most school-choosers would be from families who manage to file valid applications. Such students tend to do better in school. The described school choice system might result in a situation, in which most high-ability students leave schools that are situated in areas with a weak local student population. Schools in such areas would score badly on tests unless they manage to increase educational productivity significantly, which is likely to take some time. If there are not enough free places in better schools, many students with a low SES might then be trapped in the worst schools in the area. Such a situation would likely lead to demands for a change or abolition of the school choice system.

The last major weakness of using a CGE approach is, that a CGE contains only the hypothetical final outcome of a convergence process, which might take several hundred iterations to reach a stable state. The CGE is a useful definition for a stable state, but as there usually is no perfect prediction for the actors, most real systems do not reach a CGE instantly. In the best case, real systems converge to a rather stable state over several repetitions of the decision process, where each repetition represents a time period. And in the case of school choice, such a stable state might even not be reached, before the underlying conditions or policies are altered. The evolution of the simulated model during such a convergence process is not analyzed, often not even captured, in conventional CGE approaches. When making predictions about the effects of policy changes, it is however important to know what state to expect after a few years, and to know how the model evolves over time. A school authority might, for example, be interested in what effects a school choice system would generate within a decade, and whether it should expect an initial decrease in quality for some schools before school quality starts to increase across the board.

4.3.3.(2) Assumptions, calibration and observable characteristics

In both Epple/Romano (1998) and Nechyba (1999), the specification of the models relies strongly on ad-hoc assumptions. The specification of Fuchs draws heavily on the specification of Epple/Romano. Especially the education production functions are assumed to be Cobb-Douglas functions with only two input variables in both studies. The parameters of this Cobb-Douglas function are chosen based on an indirect approach that uses several ad-hoc
assumptions in Epple/Romano and, by extension, Fuchs. Nechyba simply assumes the
parameters ad-hoc.
In all three models there is full information. Schools can perfectly observe innate ability and
income in Epple/Romano and individual innate ability in Nechyba. Parents can observe
school quality perfectly in all three models. For the model of Nechyba, this assumption is
relatively realistic, as school quality is only intended to reflect the quality as perceived by the
parents. But in the models of Epple/Romano and Fuchs, the outcome of the education
production function is assumed to reflect the actual quality of the school. In reality however,
parents cannot easily identify the more productive schools. Moreover, all actors in both
models have preferences that do not depend on their characteristics. In reality, parental school
choice systematically differs with variations in student characteristics in a way that has
important implications for sorting. Moreover, the school choice functions of parents in all
three models are not empirically estimated and rely heavily on assumptions. In the first two of
these utility functions, distance to school is not included. This has two important implications.
First, in these models, households are footloose. They will send their children to any school
that they think to provide the best schooling. As perceived school quality is the only relevant
characteristic, within the boundaries set by tuition fees and the externally determined set of
available schools, the slightest advantage in school quality would attract all parents within
these boundaries. Second, if distances do not enter the utility function, it has no impact on the
probability of a student to attend a school if she lives close to it. Thus, the local student
population has no effect on the school, while in reality differences in most of the common
measures for school quality are driven to a large extent by variations in the characteristics of
the local students. Fuchs incorporates a spatial dimension. In his model, distances matter for
parental school choice and aggregate school characteristics.

4.3.3.(3) Modelling school effort
The driving force of the increase in overall educational productivity, that is expected to result
from school choice according to Friedman’s mechanism, is an increase in the educational
productivity of individual schools as a reaction to competitive pressure. For this mechanism to
work, schools need to be able affect their educational productivity.
But In Epple/Romano (1998), school results are determined by average peer ability only. In
Nechyba (1999), only average peer ability and spending per student determine student
achievement. And in Fuchs (2007), average peer ability, the average socio-economic status of
peers and school size affect school quality. Thus, educational school productivity cannot be
affected by school effort in all three models. This does not impinge upon the analysis conducted by the researchers in the first two models. Epple/Romano focus on the effects of vouchers on sorting and Nechyba is interested in the effects on the location decisions of households. In both models, individual and overall educational school productivity are not of immediate interest. Fuchs analyzes the impact of different types of vouchers in different school choice regimes on educational equity and on achievement. For the impact of equity, the same arguments apply as for the first two models. Concerning gains in achievement however, the exclusion of effects of different choice regimes and types of vouchers on the incentives and thus the effort and productivity of schools might be problematic for the reliability at least of the extent of achievement gains.

However, to incorporate an endogenously determined educational productivity would require models that are much more complex and it is doubtful that the potential increase in precision is worth the necessary additional effort for these three models and the questions they analyze. Moreover, there are not much insights in the literature on what determines educational school productivity. Most available measures of school quality are based on raw student outcomes, are mostly driven by differences in student intake and are therefore only a weak proxy for educational school productivity. This includes most monetary school productivity measures, which define productivity as student outcome per dollar of education funding. If the characteristics of the student intake are not taken into account, differences in this kind of productivity measure are also strongly affected by differences in student intake and are therefore no suitable measure for differences in educational quality.

Those few official quality measures that do reflect the educational quality of schools well, like the value-added measures that were recently introduced in England and for elementary schools in the CPS, have not generated a sufficient number of years of outcomes to estimate an educational school productivity function (SPF). And in order to compute such value-added functions herself, a researcher has to know student-level outcomes, the schools that were attended by students and a comprehensive set of student characteristics. To my knowledge, there is no study yet that has both good measures for educational productivity and estimates what determines this productivity.

The parameters and shape of the school production function determine how schools react to competitive pressure and have a strong direct impact on student outcomes. Changing the shape or the parameters would therefore have a strong effect on the resulting CGE. Given that there are no agreed-upon parameters, one would have to make assumptions that strongly affect outcomes and thus makes any findings vulnerable to criticism. To avoid this
vulnerability, it is possible to estimate the SPF. As institutional settings can change the incentive structure for schools considerably, the educational school productivity function should be estimated for a real school choice system. Then, this estimation should only be used to simulate a school choice system that shares as many settings as possible with the one that provided the estimates. These estimations are laborious and require a real school choice system that is both similar to the one represented in the model and provides sufficient data.

4.3.4. Implications and how to proceed with research on the effect of institutions and regulation in school choice systems

Settings in several institutions and regulations can strongly alter the outcomes of school choice systems. In order to gain further insights into the effects of specific school choice systems, it is therefore necessary to learn more about the effects of settings in individual regulations and institutions. The effects of settings in individual regulations are often affected by settings in other regulations however. Moreover, the effects of one setting can also be altered by varying external conditions, such as the geographical density of schools. A third group of mechanisms, that have a strong impact on outcomes of school choice systems, are interdependencies of decisions.

If a model shall be capable of identifying the effects of settings in individual regulations or institutions, it has to include both the most important mechanisms that are at work in school choice systems and the external conditions that have a strong effect on outcomes. If any of the important mechanisms or conditions is excluded, any findings are likely to be distorted. Unfortunately, it is not possible to include all these mechanisms and conditions in traditional approaches. Both empirical and theoretical approaches have difficulties in dealing with heterogeneous actors, the interdependence of decisions and varying institutional settings. Empirical approaches additionally suffer in most school choice systems from a lack of data on many important factors and have a hard time to deal with additional problems such as reverse causality, delayed effects and the impact of varying external choice options. Theoretical approaches are not capable of representing heterogeneous characteristics and preferences of actors. If they try to do so by using several representative agents, models become intractable. Computational models were proposed as a way forward, for example by Caroline Hoxby (2003a). By using the tremendous computational capabilities of modern computers, these models can easily incorporate heterogeneous actors and the interdependence of decisions, while offering at the same time fully controlled conditions. This opens the possibility of ceterus paribus changes to individual institutional settings. First applications of such
computational models, in the form of computable general equilibrium models described above, to school choice were capable to gain insights into the effects of some settings of school choice systems, like the denomination of vouchers. These insights were gained from models that were too complex to yield results with non-computational methods. But in order to be capable of identifying the effects of most other institutional settings, a model is needed that is more complete. Especially the mechanism by which schools can improve educational productivity as a reaction to competitive pressure is important. Moreover, to insulate against effects of variations in settings and conditions (density of schools, student population etc.) it is best to base a model on a real school choice setting and to calibrate this model with estimates that were derived from the same school choice system. And if not only a hypothetical final state that might only be reached after decades is of interest, then a deviation from the CGE approach to a model that simulates the changes from year to year is necessary.

To develop such a more complete model, and to demonstrate its application by analyzing the effect of a change in one institutional setting on school choice outcomes, constitutes the remainder of this dissertation. I will present the approach that I have developed in the next chapter.
5. Simulation approach to analyzing alternative institutional settings

I will begin this chapter with an outline of the basic strategy pursued in my simulation approach. Against this backdrop I will point out the features that characterize my simulation model as compared to those that were used in the previous literature. And I will explain why I have chosen to implement exactly these features (5.1.). Then I will present in detail the approach that I have developed to analyze the effects of alternate institutional settings in school choice systems (5.2.)

5.1. Simulation strategy

The basic idea is to create a simulation environment in which the interdependence of decisions that exists in real school choice systems is fully developed. Only with such a rich simulation model can one hope to capture the dynamics of the portrayed system that unfold over time. With these dynamics properly simulated, it will then be possible to explore the effects of changes in individual institutional settings.

This approach follows the pioneering work of Epple/Romano (1998 and 2004), Nechyba (1999) and Fuchs (2007) described in section 4.3.2. In my simulation model I have however included several refinements and additions. In the remainder of this section I will describe these changes and my motivation for implementing them.

5.1.1. Agent-based modelling versus computable general equilibrium modelling

Unlike Epple/Romano and Nechyba, I do not set up a full computable general equilibrium (CGE) model. In the CGE approach, a simulation model is run through repeated iterations until it has attained a relatively stable state for a fixed set of actors. School authorities who consider changing institutional settings are, however, not primarily interested in a distant stable state for a fixed population. School authorities are more interested in the effects that changes in institutional settings will cause in the first few years for those student cohorts who will enter their schools during these years.

Therefore, I am not primarily interested in how a relatively stable state that might - or might not - be attained in a distant future, is affected by changes in institutional settings. My focus is rather the effects in the first few years, how these effects change during these first years, and the channels through which the change in policy affects the behaviour of individual actors.
This type of micro-simulation, where the behaviour of individual agents is simulated, is called agent-based modelling. The CGE is one definition for an equilibrium of an agent-based model.

The following questions illustrate the focus of my interest: (A) Which types of students choose schools? (B) How frequent is active school choice? (C) How do choosers benefit from choice, what happens to non-choosers? (D) How do the answers to B and C differ across student types? (E) What are overall effects on sorting and academic achievement? (F) Is there a convergence to any steady state and if so, how fast is this convergence? Or will the model show fluctuations without a clear development towards a steady state?

5.1.2. Effect of competitive pressure on productivity

The main channel through which school choice is commonly supposed to increase the academic quality of schools is the following: the possibility to loose students to other schools generates competitive pressure on schools, which react to this pressure by increasing effort.

But neither of the pioneering authors has included this channel in his model. In Epple/Romano (1998) the mean ability of students at a school is the sole determinant of school quality. In Epple/Romano (2004) and in Nechyba (1998), school quality is affected only by mean ability and expenditure per student.

I have included an instrument which allows schools to change their educational productivity. In my model, schools choose the level of effort that they invest in educating their students based on the competitive pressure from other schools.

5.1.3. Imperfect information of parents about school quality

I model the information that parents have about the quality of individual schools. In the previous literature, the authors assumed that parents and other schools are fully informed about the educational productivity of each individual school. But the information that parents have about school quality in reality is usually based on official quality measures. If student characteristics are not accounted for, these official quality measures are strongly driven by differences in the characteristics of the student intake, and are only weakly correlated with educational productivity. Variations in available information have significant effects on which students apply to which schools and thus on the outcomes of school choice (see 3.1.2.(1)). In my model, parents and schools observe official quality measures that are based on the test outcomes of previous cohorts of students. As in real school systems, these tests are taken years after the students entered a school. The quality measure that is computed based on these
tests will therefore reflect the quality that an older cohort has experienced during its time at
the school. This is not necessarily the same quality that a new cohort of students, who now
enter the school, will experience. The fact that quality measures are computed based on tests
that are taken years after school entry means, that there will be a time lag between a change in
educational quality and the period when the effects of this change can be observed by parents
and other schools.

5.1.4. Fixed locations of households and schools

I assume fixed locations of households and schools. This assumption allows the local student
population to affect the student composition of schools. This composition has been shown to
have a strong impact on the quality of schools according to commonly used measures for
school quality. Fixed locations also allow the locally varying density of students and schools
to have an impact on the schooling choices that are available to students and on the
competitive pressure faced by individual schools. I will elaborate more on the motivations for
fixed locations and their effects in 6.2.1.(3).

The assumption of fixed home locations is a difference to previous CGE-approaches to
analyzing school choice. In the papers by Epple and Romano, there is no spatial dimension
and students can attend all schools without incurring commuting costs. In the papers by
Nechyba, each household is initially endowed with a house and the location of the house
determines that available public school. It would thus be necessary to buy another house in
order to change the student to another public school. In the model of Fuchs, the locations of
schools and student homes are fixed, as they are in my model.

5.1.5. Extensively calibrated model

I design and calibrate my model using real world data, institutions and actors. Using data
from the Chicago Public School System (CPS), I estimate the three central equations: the
education production function, the utility function underlying parental school choice and the
function that determines school productivity based on competitive pressure. Institutional
settings, e.g. how applications and admissions to schools are handled and which schools are
open to choice, are taken from the CPS. The students are real students who entered high
school in Chicago in the first years of the millennium. For these students I have
comprehensive information on home location and personal characteristics such as ethnicity,
gender and previous academic achievements. Schools are real public schools in Chicago. For
these schools I have information on location, an initial student body composition, assignment
areas whose inhabitants attend the school unless they successfully apply to another school and starting quality measures.

Thus, while designing and calibrating the model, I will not have to rely on ad hoc assumptions nor will I have to rely on evidence from several different sources. The uniform source of information is important because of the interaction effects of conditions and institutional settings. Take, for example, observations about the propensity to apply to other schools that is estimated for the school choice setting of a district in the suburbs, where there is scarce public information about academic quality. This finding might not, without further ado, be used in an artificial model that has diverging settings, for example about publicly available test results of all schools. Nor could it be used in a model with a different population. Parents with a lower SES for example, are on average less informed about academic quality and their school choice decisions are therefore driven less by this factor. Simply to transfer school choice preferences found for a student population in one setting (like suburbia) to another (like in an inner city) is therefore not advisable.

In my simulation model, all conditions and institutional settings are therefore taken from one school choice system. Moreover, the model holds the same population as the school choice system from which I have taken the institutional settings and conditions. And the model will be calibrated with data that were generated by this same population that was faced with these same settings. Thus, the entire school choice system in my model will be all of a piece.

5.1.6. Initial conditions taken from real world data

Another advantage of using real world data is that these provide plausible initial conditions for schools concerning average student characteristics and quality measures. The spatial distribution of student characteristics is the sum of location choices made by households. These location choices depend, besides school quality considerations, on factors such as housing prices, the distribution of employment possibilities, the closeness to relatives and infrastructure. Thus, the distribution of student characteristics is always only partially determined by school characteristics. In a school system like the CPS, with free choice among public schools and a high density of these schools, parents have a high degree of choice from most possible household locations. Under such circumstances, school characteristics should have a rather low impact on the distribution of student characteristics. Moreover, there is a high share of low-income households in the CPS, who can afford only a limited choice of housing, which also constrains location choices.
As parents are adverse to long commutes for their children, distance matters in parental school choice decisions and most students attend a school that is rather close to their home. Thus, the distribution of households affects the student composition of schools, even if there is free school choice. This student composition affects, in turn, the academic outcomes for schools according to several types of quality measures.

The spatial distribution of schools should strongly depend on the distribution of students. But as school buildings cannot be moved, the location of schools will also depend on student densities during the time when the schools were built.

Summing up, the distribution of schools and students at any point in time is likely to be correlated with variations in student characteristics. In the vicinity of a manufacturing district for example, it is likely to find numerous households which receive their income from the factories there. Most of these parents will be workers with a rather low income and limited education. Their children are less likely to be in the upper end of the spectrum of academic achievement. And they are less likely to apply to schools. This would create an environment in which schools could attract a sufficient supply of students despite providing rather low quality education. A student population that is simply generated randomly would likely not show agglomerations of student characteristics such as described above and would thus likely miss important mechanisms such as the one described above and therefore miss insights.

Moreover, the initial conditions can have an impact on the development of a simulated model over time. If, for example, average test results are used to assess school quality, schools in areas with an academically weak student population will have low average test results, will thus appear weak initially and will therefore not attract many students. If parents can only observe average test results, a school that starts with a weak quality rating cannot attract good students, will therefore have a weak quality rating again, can again not attract good students, and so on. Such a school might be trapped in a circle of weak quality ratings and academically weak students for a long time, all due to the characteristics of the initial student composition. Thus, the initial conditions like school and student density, student characteristics and their distribution over residential areas and attended schools are likely to have an impact on the development of a simulated model over time.

Moreover, with randomly generated actors and schools it is difficult to identify plausible starting values for school characteristics. The quality that parents expect from individual schools is usually based on quality measures of the school for previous years. And the level of educational productivity of a school in the first turn could be measured given data on the academic achievement of students. In the case of artificially generated actors, these values
would have to be chosen, however. And starting a simulation with the wrong values can induce spurious results that shroud the very dynamics which the model is intended to analyze. One example is a starting value for average educational school productivity that is set too low. Starting from a low level, it is not hard for schools to increase educational productivity. Thus most schools will do so if they are under competitive pressure. This increase in individual educational productivity will likely improve quality measures across the board, but more so for schools that face a higher competitive pressure. Thus, both absolute quality measures and differences in quality ratings will increase in the first years of the simulation. This development will not be a consequence of the change in the school choice policy that I want to analyze, but simply a consequence of having assumed an unrealistically low educational productivity at the beginning. And this development will unfold on top of any real effects of school choice and thus distort or hide these effects.

By using real world data to calibrate the model and to estimate starting values, it is possible to circumvent the problems that I have described above. Therefore, I use the state of the CPS in September 2004 as my starting point. This state includes the student population, characteristics and locations of schools. It also includes the institutional settings and estimates of the education production function and of the decision functions of schools and parents. Based on this model I will assume changes in the institutional settings and run the simulation model for the following years.

I want to note here, that it might theoretically be possible to randomly create a distribution of students and schools that is realistic enough to circumvent the problems that I have presented above. But to achieve such a realistic distribution, it would at least be necessary to gather information on the spatial distribution of schools and of students and their characteristics such as ethnicity, SES and innate ability in real school systems. And it would be necessary to gather information about clustering according to these characteristics. Then it would be necessary to generate schools and a student population that show these “stylized” characteristics. Only then would it be possible to analyze the effects of alternative institutional settings for a district that shows a similar student body. A random generation of actors means an additional source for possible errors however. And if detailed data on individual students, their locations, their characteristics and the schools they attend are available, and it is intended to analyze the same school system that generated this data, it would therefore be preferable to generate actors directly from the data than to use random generation techniques.
5.2. Analyzing alternative institutional settings

To analyze the effect of alternative institutional settings on the outcomes of school choice, I will proceed as follows: I start by setting up a model in which the institutions, population and calibration are all taken from one real school system, the CPS. I have chosen the CPS as the source for my model, because this school choice system is in most respects almost ideally suited for school choice to develop its full potential. Applications to schools are easily done in the CPS, students are free to apply to any and as many schools as they like, and most slots are free to all students. Moreover, most schools are not allowed to pick among their applicants those student types that they prefer, and they face strong incentives to attract many applications. The CPS also is an inner-city school district with a dense population and a high density of schools, offering several schools within commuting distance of all households.

In order to generate a simulation of a school choice system that is all of a piece, it is advantageous to base this simulation on a real school choice system that is well documented. In this respect, the CPS is almost ideal again. Most institutional settings are well documented, which is not usual for school choice systems. Some additional details I could learn from the helpful staff of the Department of Data Management at the Office of Research, Evaluation and Accountability of the CPS. The school district also keeps a vast array of data on individual schools and students that it is willing to share. Moreover, I can draw on several high quality studies\footnote{These include Lauen (2007a and 2007b), Cullen/Jacob (2007), Cullen et al. (2000, 2003, 2006), Jacob (2002) and Ferreyra (2007).} that are based on CPS data. In particular, these studies allow me to verify and supplement the findings of my own calibration estimations as well as the information I have found about institutional settings in this particular school choice system.

Once the model is thus set up - geared for strong effects of school choice and created all of a piece - it is possible to alter institutional settings one by one, holding all other institutional settings fixed as well as the population and the calibrated equations. It is necessary to be careful here though. Some settings can be altered without further ado. One example for a setting that can be altered quite easily is the type of school quality measures that are made publicly available. Most parents are unlikely to have an in-depth understanding of the mathematics that generate the measures that they see in the newspapers or on school report cards. And parents mainly care about the relative academic quality of the schools that they
consider sending their children to. This kind of comparative information can easily be deduced from ranking places of schools or from their outcomes relative to those of other close-by schools, without a deeper understanding of the mathematics behind these numbers. It is therefore likely, that the reaction of parents to quality measures does not change strongly, if these measures are computed differently, if the new measures are presented in a familiar way. One possibility for such a change in the computation of quality measures is the change from the percentage of students scoring above a given threshold to the mean test result of students. It is fairly safe to assume that parents will react to differences in ranking places or quality measure values that result from either method of aggregation in roughly the same way. At least as long as the mean and variance of the values that are computed by the new method are on a rather similar scale as the old quality measure.

But other changes in institutional settings might change the choice system in such a fundamental way that it is no longer safe to assume that this change happens ceteris paribus. One example for such a change is a restriction of the share of students who are allowed to leave their assigned school. The students who are assigned to a particularly bad school might all want to leave this school, but if only a small share is allowed to leave, then the school has much less to fear from competition, than if all students could leave to other schools. Thus, a restriction of the share of students who are allowed to leave the assigned school would reduce the competitive pressure that is generated by other schools. To account for the weakened effects of competition, the decision function that determines the educational productivity of schools would need to be adapted. It would therefore be necessary to assume a changed decision function for schools. Assumptions about a crucial part of the model, like a decision function, have a strong impact on simulation outcomes. Such assumptions have to be made very carefully, and to miss the necessity for adapting a decision function would badly damage the credibility of the model. It is therefore necessary to carefully think about whether a change in an individual institutional setting represents a ceteris paribus change or whether it is necessary to adjust the decision functions of any group of actors.

Once an institutional setting is identified that can be altered without effects on the decision functions, or once assumptions are made for the necessary adjustment of decision functions, it is possible to analyze the effects of changes in this setting. To make this analysis, several versions of the model are run, one version for each variation of the setting. By analyzing the outcomes of the different versions, it is possible to analyze the effects of the underlying change in the institutional setting. As a first application I will analyze alternative types of information about school quality.
It is important to note that this approach is not immediately capable of predicting the effects of an introduction of school choice. The reason is, that the introduction of choice fundamentally alters incentives for the actors and therefore alters their decision functions. One example for such a change in incentives is that in a school system without choice, schools face much weaker incentives to react to the quality of other nearby schools than in the same school system with school choice. Moreover, if there is no choice in the school system, then it is impossible to estimate the parental school choice function. If one wanted to use this approach to predict the effects of an introduction of school choice, it would therefore be necessary to make crucial assumptions for all the decision functions.

I have developed my approach to compare the likely effect of changing between different alternative settings of one institutional feature for a given school choice system, while holding its population and all other institutional settings fixed. Such ceteris paribus conditions make it possible to conduct the analysis without, or at least with only minimal reliance on assumptions. When this kind of analysis has been performed several times for the same institutional feature in several different school choice systems, it should be possible to generalize the effects of alternative settings for this feature to other school choice systems. But to make reasonably precise predictions of the effects of a change in this one institutional setting, it will still be necessary to consider the respective circumstances. And we will only be able to predict the effects of introducing school choice in an individual school system, once we have gained a fairly good understanding of the effects of all those institutional settings that are likely to have a strong impact on outcomes.

25 It might be possible to get some insights into parental choice preferences by analyzing existing choice options outside the public school system. If there are private schools or charter schools, choice behaviour among those schools might give an indication of the shape of the parental decision function regarding school choice. However, parents that send their children to private or charter schools are unlikely to be representative for all the parents in the district. It is not possible to generalize an estimated decision function for those parents to other parents in the district, who do not make use of this type of choice. In the case of a high density of school districts, parental choice behaviour via Tiebout choice would give a better approximation for a parental choice function for all parents. In the case of a high degree of Tiebout choice, the introduction of free choice among public schools would represent a strong increase in choice options and would therefore have relatively weak effects.
6. The Model

In this chapter I will present in more detail the agent-based computational model on which the following simulations will be based. In 6.1. I present the basic set-up, so that the reader has an overview of the model. In 6.2. I then describe modelling choices in detail and explain why each choice was made in this particular way. Some modelling choices can however only be taken after the empirical analysis of the CPS data and the calibration of the model are finalized. These decisions include a) the exact shape of decision functions, b) all parameters and the selection of school characteristics that enter the education production function and c) the utility function underlying parental school choice.

6.1. Basic set-up and sequence

6.1.1. Description of the basic model set-up

The model is populated by several thousand students and several dozen schools. These numbers will allow me to simulate the situation in a large school district, where interaction effects between and among both types of actors can fully develop.

The households live for 4 periods, but are only active in one period. Each household has either one student who is about to enter high-school (in 9th grade) or one student who already attends a high-school. Students are characterised by their innate ability and individual characteristics like gender and the socio-economic background of their household. Students have fixed home locations that are spatially distributed across the area of the school district. Home location, innate ability, individual characteristics and socio-economic background of each student are taken from real world data from the CPS and are thus exogenously given.

Each prospective freshman has a home school to which she is assigned (based on household location). Each student is guaranteed a place at this home school. The households containing prospective freshmen then can apply to any number of other schools before the beginning of the school year. Students who are already attending a high-school do not change schools anymore. The decision to apply to other schools at all is separated from the decision to which schools a student applies, if she applies to any schools for reasons explained in 6.2.1.(6).

Whether a freshman applies to any school other than her assigned school depends on student characteristics such as socioeconomic background and innate ability and on characteristics of the assigned home school. As different student types have shown differing school choice
behaviour, I will differentiate parental choice functions by student type. Student types might be identified based on innate ability, income of the household, race or other student characteristics. As a differentiation along several dimensions leads to very small groups of students, I will choose that dimension, across which choice behaviour differs most.

The probability for student $I$ of type $X$ to apply to any school is:

$$P_i^X(Apply) = X_i \beta_i^p + S_{i-1}^i \gamma_i^p + \epsilon_i^p$$  \hspace{1cm} (1),

where $X_i$ is a vector of student characteristics, $S_{i-1}^i$ is a vector of characteristics for a previous cohort at the assigned school, that are observable at the time of the school choice decision. $\beta_i^p$ and $\gamma_i^p$ are vectors of coefficients for this student type and $\epsilon_i^p$ is a zero-mean normally distributed error term that represents an idiosyncratic propensity to apply to any school$^{26}$.

Whether a freshman who applies to any school applies to a particular school $J$ depends on the utility that the household expects from attending this particular school, given the assigned school. This expected utility depends upon distance to the school and on the quality and other characteristics of school $J$ and of the assigned school, as perceived by the household. Parents will apply to all schools for which the expected utility is higher than the expected utility at the assigned school.

The parental utility function for school $J$ for student $I$ of type $X$ is:

$$U_{ij} = f(D_j) \times (X_i \beta_i^U + S_{i-1}^i \gamma_i^U + S_{i-1}^j \delta_i^U + \epsilon_{ij}^U)$$  \hspace{1cm} (2)

where $S_{i-1}^j$ is a vector of characteristics of school $J$ that are observable at the time of the school choice decision, $D_j$ is the distance from the household home to school $J$, $\beta_i^U$, $\gamma_i^U$ and $\delta_i^U$ are vectors of coefficients and $\epsilon_{ij}^U$ is a zero-mean normally distributed error term that represents an idiosyncratic preference of household $I$ for school $J$. The function that determines the impact of distance on the utility that parents expect from school $J$, $f(D_j)$ is:

---

$^{26}$ One remark about the error term in functions (1), (2), (4) and (5). The specification with an error term is correct here and will be correct in the estimations in Chapter 8. Due to the command that I will later use for predictions based on these estimated functions, there will however be no error term, and therefore no random element in the predictions of the simulation as used in chapter 9. As I base my findings only on averages of entire simulation runs of aggregates of entire populations of schools or students, the absence of a zero-mean error term on the level of individual schools and students does not affect most outcomes. However, due to the definition of the official quality measure that is used in the CPS, the share of students who meet or exceed standards, this measure can be affected. I discuss the limited extent of this possible distortion and why I chose to accept it in 8.1.2.
\[ f(D_{ij}) = B + \frac{(d_{ij}^{\text{max}} - D_{ij})}{d_{ij}^{\text{max}}} \]  \hspace{1cm} (3)

where \( B \) is a fixed value and \( d_{ij}^{\text{max}} \) is the maximal distance to a school that a household is willing to accept for any set of school characteristics. \( f(D_{ij}) \) is equal to \( B \) for \( D_{ij} = d_{ij}^{\text{max}} \) and equal to \( B + 1 \) for \( D_{ij} = 0 \). Schools for which \( D_{ij} > d_{ij}^{\text{max}} \) are too far away from the student home and are therefore not considered as a possible choice by students. The way in which the impact of distance on the utility function is modelled is a consequence of the data that I have available about student home location. I will explain this modelling choice in detail in 6.2. The fact that the assigned school is included in \( U_{ij}^{x} \) is due to the fact that I estimate school preferences as the probability of attending school \( J \) if assigned to school \( I \). I will elaborate on this approach in 8.5.

Note that this parental utility function only captures utility derived from school characteristics. Based on the conditions in the CPS I assume that public schools do not charge tuition, transportation to school is free and school taxes are fixed for the single district that encompasses all schools in the CPS. I also assume that home location is not determined by school characteristics. These assumptions allow me to exclude from my model the political decision about the school tax level. And these assumptions allow me to separate the school choice decision from the choice of the home location. I can therefore exclude from my analysis the impact of housing conditions on the utility derived from school choice. Moreover, I do not have to simulate the home location decision and a rental market and/or housing market. And, as a consequence of this separation of the decision about school choice and about the home location and as a consequence of the absence of tuition and transportation costs, school choice has no effect on disposable income. Therefore, I do not have to consider consumption in the household utility function.

The measures for the quality of schools are based on a centrally administered test taken by the students attending the 11th grade. I model the individual student outcome in this test as a function of the innate ability of a student, individual student characteristics, and average innate ability of the peers and the educational productivity of the attended school. The outcome for student \( I \) of type \( X \) at school \( A \) is:
\[ \text{Outcome}_{it} = X_i \beta_s^O + S_a \mu_s^O + Q_o \mu_s^O + \varepsilon_i^O (4) \]

where \( S_a \) is a vector of characteristics of the attended school \( A \), which includes the average innate ability of the peers, \( Q_o \) is the educational productivity of school \( A \), \( \mu_s^O \) is a coefficient and \( \varepsilon_i^O \) is a mean-zero distributed error term representing an idiosyncratic effect on the outcome of student \( I \) in this particular test.

Schools have fixed locations and are spatially distributed across the area of the school district. Each student is allowed to apply to all schools. Each school is obliged to admit all students that are assigned to this school and choose to stay at the school. Once all students who stay are accommodated, remaining free places are assigned to applicants independently of their characteristics and independently of the current student composition at the school. If a school receives more applications than it has free places, students are admitted by a lottery that is held among all applicants. In this lottery, all applicants have the same probability of being admitted.

Financing follows students. The funds of an individual school are therefore proportional to the number of attending students. Facilities are provided to schools by the district free of cost. Funds are thus used for staff only, at a wage rate that is fixed for all public schools in the district. Thus, if a school looses students, it has to reduce staff. Moreover, a school that does not have any freshmen in a year is closed down and reconstituted under new management. To avoid computational complications, I assume that students stay at a closed-down school and are taken over by the new management. And I assume that the new management starts immediately with the next school year. This way, all schools are always open to admit students, and I do not need to reallocate students.

A school that has no students for a cohort cannot generate a quality measure or a dropout rate for this cohort. An empty school can also not generate average student characteristics such as the share of students from low-income households. I assume that parents and other schools expect school quality and average student characteristics of a reconstituted school to be equal to the averages for the entire school district.

Schools can choose the level of educational productivity by choosing the level of effort that they exert to educate their students. This choice is based on external pressure to achieve good results. This external pressure is induced by competitive pressure generated by other schools that vie for the same students, by the share of places of the total student capacity that remained empty in the previous year and by probationary status.
\[ P_j = F_j^{t-1} \theta^V + (F_j^{t-1})^2 \theta^W + C_j^{t-1} \eta^V + (C_j^{t-1})^2 \eta^W + p_j^t + \epsilon_j^Q \]  

(5)

Where \( F_j^{t-1} \) is the free capacity of the school in the last period, i.e. the share of the total capacity that could not be filled with freshmen in the last year. \( C_j^{t-1} \) is the level of competition that school \( j \) faced from other nearby schools, based on the newest observable set of quality measures. \( p_j^t \) is a dummy variable that indicates whether the school is currently under probation and \( \epsilon_j^Q \) is a zero-mean normally distributed error term.

The level of competition is a function which is increasing in the number of schools that are close by, the closeness of these schools and their quality relative to the quality of the own school:

\[ C_j^{t-1} = \sum_{n \neq j} f(D_{nj}, Q_n^{t-1}) \]

(6)

where \( f(D_{nj}, Q_n^{t-1}) \) represents the competitive threat to school \( J \) originating from one competitor, school \( N \). \( D_{nj} \) is the distance between the two schools, \( Q_n^{t-1} \) represents the quality of school \( N \) as perceived by the parents. This quality is determined by the newest available official quality measure. The exact shape of functions (5) and (6) will be determined based on the real choice behaviour that I observe for the students in the CPS in chapter 8.

6.1.2. Sequence of actions

The following sequence of actions takes place in each year of the simulation.

Step 1: A new generation of households enters the model.

Step 2: Agents make their decisions:
Households containing freshmen observe the school characteristics, based on the newest available data, and choose whether to apply at all and if so, to which schools to apply.
Schools simultaneously observe the level of competition that they face and choose the level of educational productivity for educating this cohort of students.
Note that in this setup no actor knows about the simultaneous decisions of all other actors when deciding about his own actions.

Step 3: Assignment of students to schools:
Students who do not apply to any other school are assigned a place at their assigned home school. Students who have applied to schools are admitted at these schools if these have a sufficient number of free places. For the oversubscribed schools, lotteries are held to assign free places to applicants.

Step 4: Computation of outcomes and average characteristics:
Test results for individual students attending grade 11 are calculated using the education production function and the level of educational productivity that was chosen by the schools for this cohort. Based on individual student test results, the official quality measure is computed. Also, average characteristics for the current student population, like the share of low-income students, are computed. These average characteristics and the quality measure will be used to inform following cohorts of households and the decisions of schools in later periods. Additionally, outcome characteristics, for example concerning sorting, are computed in step 4.

6.2. Modelling choices

In this section I will describe modelling choices in detail and also explain why I have taken these decisions.

6.2.1. Choices regarding households

6.2.1.(1) The timing of school choice
There are several motivations that can affect school choice. One is the desire of the student or her parents to get access to higher academic quality. But students also change schools if they are expelled from the school that they are currently attending or if their guardians move, which happens frequently, and often involuntarily, in US inner cities (see 3.4.3.(2)). School changes of the latter two types are only partially motivated by school characteristics. At least leaving the old school is in these cases not motivated by the quality of this school. In some cases, such school changes are not motivated by school quality considerations at all. This is the case if the household has to move and parents do not consider an application to any school which means that the student will automatically attend the assigned home school that corresponds to the new address. And if a student is expelled, she can only attend those schools that are willing to admit her. It is not possible to get more than fragmentary information on these types of school changes. For an expelled student, I do not know which schools would have accepted her. Information about household moves will also be patchy at best. Some
information might be gleaned from the assigned home school, at least whether the family home moves to the assignment area of another school. But I do not know whether a student who attends her home school after a move does so because the parents favour that school, because they do not have the time to get informed about other schools in the area or because they do not care about school quality. Moreover, popular schools will have no remaining free places once the school year has started. Additionally, school changes that take place after the high school career has started are likely to have detrimental effects as discussed in 3.4.2.(4). For this reason, most parents will be reluctant to apply for school changes, if they do not have to, once their child has started high school.

Thus, school changes that happen after the students have started high-school are much more likely to be driven by other reasons than school characteristics, such as household moves due to the loss of a job of a parent or because the family was expelled from rented housing. Such changes will also be made by a special selection of students and these will face restrictions due to the timing and/or reason of their school change.

Before entering high-school on the other hand, all aspiring freshmen have to make a school change at the same time, from their middle- or elementary school to a high school. This mandatory school choice will have an important effect on their future career opportunities. Also, these students and their parents are then surrounded by peers that have to make the same choice at the same time. And they are usually offered easily accessible information about individual schools. Thus, all parents choose schools at the same time and their choices will likely be driven by the characteristics of the schools. Moreover, all students have the same chances to get into any school if they apply before the beginning of their freshmen year.

To capture school choice decisions that are motivated mainly by schooling considerations, I focus therefore on the first choice of a high school for prospective freshmen. As each student enters high school only once and I focus on this first year, households only choose schools in one period in my model. Students exist for additional periods in my model only in order to generate quality measures and student composition characteristics of the schools they are attending.

6.2.1.(2) Explanatory variables

As described in 3.1.2.(1), parental choice of schools heavily differs across student types. Students whose parents have a higher SES have for example been shown to be more likely to apply for a school change and their applications will be driven more strongly by academic considerations. Therefore, I will differentiate the school choice functions by student type. The
definition of student types might be based, given the available data, on income status, race or innate ability. In chapter 8.2 I will choose that student characteristic to define the student types, across which the observed choice behaviour in the CPS differed most strongly.

Innate ability is a strong driver of academic success. The abilities and the previous knowledge that a student already has acquired when she enters high-school have a strong impact on academic success during high school. I approximate academic innate ability with previous academic success, as customary in the literature (see for example Wilson/Piebalga 2008, Lauen 2007a, Cullen/Jacob 2007).

Concerning other characteristics of students and schools that affect either school choice or educational outcomes, I will only provide lists here of those variables that I will use in the simulation. This choice of variables is mostly determined by the availability of real world data on those variables, on the possibility to generate those variables in the simulation and on difficulties caused by high correlations between potential explanatory variables. I will elaborate on the choice of variables in Chapter 8, once I have presented the dataset.

Student characteristics are innate ability, gender, a dummy variable on household income, race\textsuperscript{27}, a dummy variable on English language proficiency, special education status, whether the student has repeated at least one grade in elementary school and whether the student has actively chosen her elementary school.

School characteristics are always based on the newest observable data at the time of school choice. These characteristics include the official quality measure of the school, the share of low-income students, the share of students that dropped out, the share of the own race at the school and the share of students with limited English proficiency.

The simulated educational achievement of a student (function 4) is affected by student characteristics, the average innate ability of her peers and the level of educational productivity of the attended high-school.

\textsuperscript{27} Race does not enter the education production function or the choice functions. But I will differentiate these functions by race. The reason is that both school choice behaviour and the education production function in the CPS differ most strongly across race (as compared to a differentiation across income type and innate ability) and I therefore choose to identify student types by race. This choice of the variable across which I differentiate student types will be described in detail in chapter 8.2. I do not assume that this difference in choice behaviour and educational attainment is based on genetic differences. Instead, I use race as a proxy for other unobserved variables. I do not have information on the educational background of parents and on their migration status and I only have vague information on household income. I also do not know about the household situation, especially whether the student lives in a single-mother household. Additionally, the involvement of parents in school choice might be affected by the gains that parents expect from schooling which might be affected in turn by racial discrimination previously experienced by the parents. Educational background, migration status, expected discrimination and income levels of parents differ across race, as does the share of single-mother households. I therefore interpret race as a proxy for differences in parental SES and in expected gains from education.
Whether a student applies to any school (function 1) depends on student characteristics and school characteristics of the assigned school.

How attractive school $J$ is to an individual student (function 2) depends on student characteristics, school characteristics of the assigned school and school characteristics of school $J$.

6.2.1.(3) Fixed household locations and distances

One of the most important drivers of parental school choice is the distance of the schools to the family home (see 3.4.2.(3)). In order to compute these distances, I need information on the locations of households and schools.

In the US, changes in household location have been shown to be affected by school characteristics (see for example Jacob/Lefgren 2005). In the absence of school choice, parents use such Tiebout choice to get access to better schooling for their children. Contrary to the setting in which Tiebout choice usually takes place, I analyze a large school district with free school choice and many schools within commuting distance of all households. If access to schools is not tied to residence and there exists a high density of schools, then it is not necessary to move close to an individual school in order to get one’s child admitted to this school. Thus, household location and household moves are less likely to be motivated by school characteristics in such a setting. Other factors such as the location of the occupation of the parents, housing costs, crime levels and closeness to friends and family, should be more important in determining household location in this setting.

Parents might be willing to move in order to get access to better schools for their children. But in a school system such as the CPS, most parents will not be willing to face the immediate and long-term costs of a move, merely to shorten the commute of one of their children to the favourite school, maybe at the cost of a longer commute for other children. That most households do not move in order to get access to a particular school for their children can be seen from the fact that about half the students in the CPS attend schools other than their home school, as can be seen from the data presented in the next chapter. Therefore I assume household location to be independent of school characteristics and hold it fixed\textsuperscript{28}.

Moreover, there exist several modelling reason to include fixed locations in the model and to include the distance to schools in parental choice decisions.

\textsuperscript{28} Note that I am only analyzing the first choice of a high school for entering freshmen. Thus I assume the household location to be fixed only for the time between applications and the beginning of the school year.
With fixed locations and parents caring about distance to school, many applicants to an individual school will come from the vicinity of the school. This renders it possible for the characteristics of the local student population to affect school outcomes. Such an effect of the characteristics of the local student population has been shown in reality to be an important driver of differences in average test outcomes, as presented in 3.4.2.(2).

If distances are not included in household utility, all parents in the model would want to send their child to the best school according to the newest observable quality measures and other school characteristics that they deem important. They would be willing to send their child to a school anywhere in the district in order to gain the slightest improvement in the expected utility level. This would lead to vast and sudden changes in application numbers due to minor changes in quality measures each year. The school choice system would therefore be highly unstable.

If distances affect parental utility but household locations are not fixed and moves are costless, all households would want to move close to the best schools. If moves are costly, I would have to estimate these costs. Moreover I would in both cases have to include the set of available housing and some sort of pricing mechanism for housing. As households do not base their home location solely on schools and housing costs, as argued above, I would have to include other characteristics of housing that are not independent of school characteristics. The inclusion of housing decisions would thus make the simulation model much more complex. Such an increase is not justified by the small impact that is to be expected of school characteristics on household location in the face of free school choice and a high density of schools.

6.2.1.(4) The impact of distance

The computation of the distance between student home location and the locations of the school in the district is not straightforward. The reason is that the CPS officials are not allowed to reveal the home location of individual students. What I can observe is the name and administrative identification number of the assigned home school. I identified the addresses of all those high-schools that enter my simulation and then the latitude and longitude of these addresses. This gives me a numerical location that can be used to compute the distance between two schools. Students are assigned to home schools in their neighbourhood, so that the home location has to be quite close to the assigned school. The location of the assigned school can therefore be used as a proxy for the student home location.
The home school location is an imperfect proxy however for two reasons. Firstly, given the size and shape of the assignment areas, the proxy is imprecise, as the family home can be situated across the street from the school or up to a few kilometres distant. Most students live within a 1 to 1.5 kilometre radius around their assigned school, but the distance can be up to 4 or 5 kilometres. It is possible that another school, in some cases even another assignment area school, is closer to the student’s home than the assigned school (see the graphs named “Assignment Areas...” in “Daten\Stata CPS Data\Graphs and Tables” on the accompanying DVD for maps of assignment areas of the CPS). The second imperfection of this proxy is that the distance to the assigned school is, per definition, always equal to zero. As a consequence, I cannot use the distance to schools in order to estimate whether a student applies to schools at all (function 1). The reason is, that distance between the home location proxy and the location of the attended school is for non-choosers, and only for non-choosers, always equal to zero. A distance of zero is therefore a perfect predictor of not applying to any school. Such a perfect predictor makes it impossible to conduct binary outcome estimations. I could therefore only estimate the impact of distance on the school choice decision for those students who do not attend the assigned school (function 2). But this estimation would still be hampered by the fact that the distance measure is imprecise. Assume that the distance between the assigned school and alternative school A is two kilometres. If the real home location is situated, from the point of view of school A, halfway between school A and the assigned school, the real distance between home location and school A is one kilometre, and thus smaller than the approximated distance, which is two kilometres, the distance between the location of school A and the location of the assigned school. School A might even be closer to the student home than the assigned school. In that case, attending school A would shorten the daily commute as compared to attending the assigned school. If, on the other hand, the student home location is situated - again from the point of view of school A – one kilometre behind the assigned school, then the real distance is three kilometres and thus bigger than the approximated distance. This difference between real and approximated distance does not have a strong effect for schools that are far away from the assigned school. Whether the real distance is 22 or 24 kilometres does not have a strong impact on the school choice decision. But for short distances, this imprecision is a major problem. Whether the real distance is 1 or 3 kilometres should affect the choice decision significantly. And whether an alternative school is actually 500 meters closer than the assigned school or, as according to the approximate measure, 1500 meters distant, should have a strong impact on the school choice decision. In the CPS, most school-choosers travel not very far. 52% of school-choosers travel less than 6 kilometres,
82% travel less than 10 kilometres (see the do-file “C:\Daten\Stata CPS Data\Explanations & Checks\School choice behaviour in reality.do). For these distances, the imprecision should have significant effects.

An additional difficulty of estimating the impact of distance on the school choice decision is, that the impact differs systematically with student characteristics. Parents with a high SES have been shown, for example, to care less about the distance to schools (see 3.1.2.(1)).

Due to the described difficulties of estimating the impact of distance on school choice behaviour, I have decided not to estimate it along with the impact of other characteristics. Instead, I will model the impact of distance separately and calibrate this model so that it generates a distribution of distances to attended schools that matches the distribution that can be observed in the CPS. The functions (2) and (3) represent this modelling of the impact of distance on the attractiveness of schools. The multiplicative nature of $f(D_y)$ means, that the attractiveness of a school that is based on all the characteristics of a school -except the distance- is discounted by a distance factor. The form of $f(D_y)$ is chosen, so that $f(D_y)$ is equal to $B$ for $D_y = d_{\text{max}}$ and equal to $B + 1$ for $D_y = 0$. The constant $B$ prevents $f(D_y)$ from varying between 0 and 1. This is important, as trial runs have shown that in this case the school choice decision would be driven almost exclusively by distance.

6.2.1.(5) School choice
As shown above, academic quality as perceived by parents and distance of a school to the home are the most important drivers of parental school choice. Therefore, I have included these variables in the parental school choice function in my model.

Other variables that have been shown to have an important impact on parental school choice are racial and socioeconomic composition of the student body, the share of the own race at the school, conduct of current students, dropout ratios and security. In Chapter 8, I will decide which of those variables for which I have access to data to include in the estimations, based on the variables used by previous studies, the reliability of the variables in my dataset and whether I could reproduce them in the following simulations.

6.2.1.(6) Separation of the choice decision
In my model I separate the decision to apply to any school from the decision to which schools a student applies. The reason for this separation is that many students, or rather their parents, do not apply to any schools. Active school choice requires effort. Even if the application procedure is not time-intensive, parents have to gather information about the potential schools
and the application process, have to fill out papers and have to keep deadlines. To start with, parents have to care enough about schooling that they are willing to invest any time and effort into applications.

The fact that some parents do not care much about schooling, that they are not able to take the initiative for active school choice or that they balk at the required effort does not mean that they would not prefer another school than the assigned one if they were presented with a choice that requires no effort. Therefore, if I only used the second part of my simulation of school choice behaviour, that is if I only estimated the attractiveness of each school and compared expected utility levels of the assigned school to others, then all parents who prefer any school to the assigned one would be classified as choosers.

Therefore, I first estimate whether a student applies to any school at all. Then, I only compare expected utility levels at potential schools for those students who are predicted to apply to schools. Apart from preventing unrealistic choices and a much too high share of active choosers, this separation helps me reduce computing time, as the comparison of expected utility levels for each student is done in a programming-loop that takes much computation time.

6.2.2. Choices regarding schools

6.2.2.(1) Types of schools

In my simulation, I only include schools that are open to applications from all students and that are not allowed to base student admission on student characteristics. This excludes the following types of schools: private schools, selective admission schools and schools that cater exclusively to certain types of students.

There are several reasons for these exclusions. To start with, my dataset only contains data on characteristics of public schools in the CPS and on the students who attend these public schools. This makes the inclusion of private schools impossible.

Public schools in the CPS are intended to be racially integrated. School choice in the CPS was introduced in order to give minority-students access to good public schools. Therefore, there are no schools that cater specifically to ethnic groups. There was only one type of schools that catered specifically to academically weak students. The Academic Prep Centres. But these were limited to students that were not academically ready for high-school while being too old to attend an elementary school. Moreover, the Academic Prep Centres were discontinued in 2003, and are therefore only contained in the first usable cohort of my dataset. There is a number of charter schools and some of those cater to specific student groups. But these
schools are usually newly founded, very small and have restricted access. Due to these characteristics and other reason that I will explain in 7.2.2.(2), most charter schools are excluded from my analysis.

For selective admission public schools, I have only scarce information on the admission process. As I can only observe the attended schools, but cannot observe which students applied to each school, I can also not estimate the selection behaviour of schools among applicants. As a consequence, I cannot simulate the admission to selective admission public schools, except based on strong assumptions. If I would simulate this admission process, it would add at least two layers of complexity to my simulation. Firstly, I would have to simulate a selection process. Moreover, I would have to separate the admission process into two levels, one for selective admission schools and one for free admission schools, which would probably make it impossible to use a crucial simulation shortcut that is described in 6.2.3.(4). Secondly, not all schools would compete with each other on the same level. The pool of potential applicants for selective admission schools differs from the pool of potential applicants for non-selective schools.

I have therefore decided to focus on schools that are open for applications from all students and that are non-selective. This means that my student sample is reduced to those students that did not get access to private schools or to selective admission schools. These students are a sub-sample that has a slightly lower innate ability and SES than the entire student population in the CPS. But it is a student sample that is homogeneous concerning school choice possibilities and admission probabilities. Moreover, it is a sub-sample of the student population that has no schooling options besides freely-accessible public schools and therefore has the strongest need for a high educational quality of these schools. And it is the student sample for which an increase in public school quality would likely have the strongest beneficial effects regarding dropout rates. Moreover, the school sample is one of schools that compete for students on an equal footing. Or rather on an almost equal footing, as there are assignment area schools in the CPS, to which students from a given area are assigned. And there are schools without assignment areas that are only attended by students that actively applied at the school.

Although there are good reasons to exclude private schools and selective admission schools, it is important not to forget that these schools can in reality have an effect on non-selective public schools. The channel for this effect is the self-selection of students into the freely accessible public schools. If the quality of freely accessible public schools increases, it is likely that some students who would have chosen private schools or selective admission
schools before now apply to freely accessible public schools. These decisions would lead to an influx of students who are on average academically stronger than the average student in the freely accessible public schools. Such an influx of academically strong students could improve average test results directly. And the benefit of academically stronger peers could also lead to a slight improvement of those students who are in the freely accessible public schools in any case. A decrease in overall school quality would have the opposite effect. Academically strong students leave, average peer ability decreases and overall outcomes decrease. Thus, self-selection of students into and out of the outside options of private and selective admission schools could act as an amplifier for the effect on average outcomes that originated in an underlying change in academic school productivity\(^{29}\).

### 6.2.2.(2) Reactions to competitive pressure

In a school choice system in which financing follows students, as in the CPS and in my model, schools will have to reduce staff if they loose students. There are several ways to portray the impact of the threat of this potential loss of staff positions on the decisions of schools.

A straightforward approach would be to use the teacher/student ratio. A school with a high ratio will have to reduce staff if it looses students. But in order to choose the desired level of educational productivity, a school would have to gauge the expected number of lost students for every possible effort level. To achieve this end, an individual school would have to consider the quality of other nearby schools and their overall attractiveness to students. It is however unlikely that a school can accurately predict the decisions of other schools, and the reactions to the outcomes of these decisions by parents. Without this information, no school could produce a reliable estimate of the number of lost students for each level of educational productivity. Moreover, teachers and principals would not want to risk relying on such predictions if a wrong prediction implies the loss of positions. Consequently, teachers and schools have been shown to immediately react to an arising possibility of student loss. They did not wait and see how many students would actually leave their schools before changing their effort spent on educating students (see 3.2.2.(1)). Schools and teachers reacted to the threat of loosing students that could be derived from the competitive pressure that they faced from other schools.

---

\(^{29}\) This insight is based on a helpful comment at the presentation of an earlier stage of my work at the ECER 2010 conference in Vienna, by a colleague from Britain, whose name I unfortunately cannot remember.
Moreover, using the teacher/student ratio would cause discontinuity in the model. A school that has a high teacher/student ratio is at risk of losing staff positions and will thus exert high effort. If this school loses a few more students, it will have to lay off one teacher. By doing so it reduces the ratio to a value at which no further position at the school is endangered. Therefore, the school would be likely to exert less effort than before the layoff. Moreover, the effects of changes in student numbers are not linear. If, for example, a school with a high teacher/student ratio attracts a few students, the positions of all teachers will be secure. If the same school attracts a few more students, it is assigned one new teacher, the ratio rises and positions are insecure again. If the school attracts still more students, positions are secure again until a new position is assigned etc.

This effect on school effort, which arises from small changes in the number of students, also varies strongly with school size if it is based on the teacher/student ratio. The smaller the school is, the stronger the teacher/student ratio will change when one teacher is laid off. And the larger a school is, the smaller is the share of students that has to leave to make positions secure or insecure. Thus, if the student/teacher ratio were used to determine school effort in a simulation, most schools would show frequent fluctuations in effort as a result.

Summing up, the teacher/student ratio is an indicator of the immediate threat of the loss of staff positions. But in order to base the exerted effort on this ratio, schools would have to be able to make accurate predictions of the actions of other actors. They would also have to be willing to rely on these predictions. Even if this were possible, the use of the teacher/student ratio to determine effort in a simulation would lead to frequent and strong fluctuations in effort levels of most schools.

Therefore I chose to use competitive pressure to determine the effort level of schools. I do not assume that schools predict the exact number of students lost per turn for a predicted effort level for all other schools. Instead I assume that an individual school will base its effort decision on the combined potential of other nearby schools to lure students away.

To use this approach I have to sacrifice the use of the teacher/student ratio, the variable that directly determines the loss of staff positions. But in return I do not have to make unrealistic assumptions about the precision of predictions by schools and about the extent to which schools rely on such predictions. Moreover, this approach results in a consistent reaction of schools to rising levels of threat.
6.2.2.(3) Explanatory variables for external pressure

Academic quality and distance are the strongest drivers of parental school choice. Therefore, a school is more likely to lose students to another school, if this other school can demonstrate a higher academic quality. A school is also more likely to lose students to another school, if this other school is located closer to the own school. The closer a school is, the lower the increase in commuting costs for the average student in the assignment area, if this student were to choose the other school. And each additional school within a reasonable commuting distance increases the choices that are open to parents. Thus, competitive pressure increases with the number of close-by schools.

I therefore assume competitive pressure to increase with an increase in the following three factors: (1) the number of nearby schools, (2) the closeness of these schools and (3) their quality relative to the quality of the own school, as perceived by parents. This perceived quality is based on the newest available official measure for academic school quality. This academic quality can be affected by schools via a change in effort that results in a change in the educational productivity. The exact shape of the competitive pressure function will be determined in 8.7 based on the observed choice behaviour in the CPS.

Hastings et al. (2007) use a similar approach, in which the authors measure competitive pressure based on the average distance to those other schools that are located within 5 miles of a school and on average test results (the official quality measure that was available to parents) for those schools.

Another factor that might affect the effort decision of schools and teachers is the share of the total school capacity of places that remained empty in the last year. The total capacity is based on the size of the facilities, for example the number and size of available classrooms. The share of this capacity that remains free does not affect the number of staff positions as long as it does not change from year to year. But this share can affect the effort decisions of teachers in two ways. Firstly, it is a measure for the success of the school in attracting students in the previous year. It represents the realization of the potential of competing schools to attract students away from the own school in the past and can thus add to the estimation of the expected potential of other schools to attract students away in the future. Secondly, it is a measure for the threat of school closure. The CPS closes schools for underutilization, in part to reconstitute them under a new management (NCBG 2007). In this case, usually all staff positions are lost. All teachers loose their position, and have at least to apply to free teaching positions at other public schools in the CPS. The higher the share of the free capacity, the higher is the probability of a school closure.
A third factor that might put a school under external pressure to increase educational productivity is probationary status. In the CPS, a school will be declared to be under probation if it fails to meet given standards on a set of characteristics, which include student test results. A school that is under probation receives additional help. But if a school is on probationary status for an extended time, it might face more unpleasant consequences such as principal removal, restructuring or even school closure (CPS Performance Policy).

I also include the square of competitive pressure and free capacity. These squares show a strong impact in the estimations. The reasoning behind the inclusion of these squares and their significant impact is threefold. Firstly, increasing productivity is costly, most likely with diminishing returns of productivity for the spent effort. Therefore, the reaction as measured in increases of productivity to further increases in external pressure should be lower, the higher the level of external pressure is. Including the squares of the explanatory variables makes it possible to represent diminishing returns. Secondly, as an alternative to increasing educational productivity, which requires a lot of additional effort, schools could improve their outcomes through various kinds of cheating. The higher the effort-costs of a further increase of educational productivity are, the more likely it is that schools resort to these alternative means. Thirdly, teachers might give up on schools when any effort the teachers might expend is not likely to get the school back on track. If, for example, a school has a free capacity exceeding 60% of available places, it is likely to be closed soon for underutilization. Many of those teachers who can do so are likely to try and get into other schools at this point. As the best teachers will have it easier to find a position at another school, the average educational productivity is likely to decrease as a result. The teachers who remain at the school are likely to reduce their effort that they might expect to be spent on a doomed enterprise. Likewise, a school that is faced with too strong a competition from nearby schools might be considered a lost cause which might result in reduced effort and leaving teachers.

30 In a sidetrack of my research, I found preliminary evidence that schools in the CPS increase the extent of various cheating measures as a response to increased external pressure, especially by trying to avoid a high share of iep students, as these students are less likely to meet or exceed state standards. This approach was meant to be used as a safeguard if I would not manage to measure an impact of external pressure on productivity directly. As educational productivity cannot be observed by parents and schools in the CPS were judged based on outcomes only, it was not unlikely not to find a direct impact. In that case, I would have estimated the impact of external pressure on cheating by schools, and then started to model a function for educational productivity based on this reaction. As I did find a direct impact of external pressure on productivity, I discontinued the sidetrack in a state, where findings were only preliminary. The interested reader can find the latest state of this research in the folder “C:\Daten\Stata CPS Data\Data Preparation Schoolmax\School cheating”, mostly in the do-files “Find school cheating effects.do” and “IEP changes effects on schools 2.do”.
Based on the considerations above, I will use as explanatory variables for external pressure, and thus for productivity, the competitive pressure from other schools, free capacity in previous years, the square of these two variables, and probationary status.

6.2.2.(4) Decision-making process
The productivity of schools decreases with an increase in the influence that teacher unions have on working conditions of their members. And when schools or teachers were given additional incentives to exert more effort, either by carrot or by stick, they responded with an immediate increase in effort (see 3.2.2.(1)). This evidence indicates that the average school in a setting without free school choice is capable of increasing effort. The evidence also indicates, that to induce such an increase in effort it is necessary to offset the associated costs with rewards or with the possibility to prevent unpleasant consequences. Therefore I assume that an increase in effort is associated with a loss of utility for schools and teachers. Note that I do not deny that teachers are also intrinsically motivated. But I do assume that the average teacher does not like to increase effort above his intrinsically motivated level.

The loss of students increases the risk of unpleasant consequences in most school systems. Possible consequences in the CPS -and hence in my model- are relocations or even lay-offs of staff. In extreme cases, the loss of students or continued bad test results might even lead to the closure of the school. The declining reputation and the increasing pressure from school authorities that are likely to be associated with a considerable loss of students are another source of distress. Rising external pressure means that such unpleasant consequences become more likely. To decrease the probability of unpleasant consequences, schools can improve their quality measures by increasing their level of educational productivity. Improved quality measures would reduce the impact of all three components of external pressure. Improved quality measures would immediately reduce competitive pressure (1). They would also lead to more applications and consequentially to a reduction in the share of free places (2). And improved quality measures would reduce the probability of being on probation (3).

In chapter 8.8.1, I will estimate how the level of educational productivity is determined by the three sources of external pressure that I have identified above, competitive pressure, free capacity and probationary status.

Changes in the institutional settings can have strong effects on incentives for schools and teachers. A change in an institutional setting could therefore result in strong and sudden changes in the intended level of educational productivity. But educational productivity is based on school characteristics that cannot be changed at will from year to year. The average
educational quality of the teaching staff, for example, can only be changed significantly if a large share of the staff is replaced, and new teaching methods need to be tested and introduced. Therefore, I will restrict the actual yearly change in the level of educational productivity to changes that were observed in reality for the CPS.

6.2.2.(5) **Schools are unitary actors**

The school as an actor in my model consists of the school administration and the teachers. Any decisions are taken and executed by members of one or both of these groups. This includes the setting of the level of effort to be exerted in the following school-year. I have decided to treat the decision making inside schools as a black box. There are two reasons for this decision. The first reason is that modelling the internal decision-making of schools would add a layer of complexity to the model without worthwhile additional insights into the main areas of interest of this study. The second reason is that it does not matter for my analysis to know how the school staff agrees on a course of action and how it achieves a change in the quality of the education that the school provides. What is relevant for my analysis is solely whether schools intend to react to external pressure by increasing effort and whether they are capable of improving academic quality. As presented in 3.2.2.(1), schools are capable of doing so.

6.2.2.(6) **Fixed locations**

Schools are tied to facilities. These facilities are expensive to build and to maintain. A school district has therefore usually about as much physical capacities as it needs to accommodate the students who are living within its boundaries. School facilities are rather specialized. This makes it hard to either sell such facilities or to buy new buildings that do not need extensive and costly reconstruction before they can be used as schools. Moreover, schools are usually built at a location where there is a need for them, usually close to residential areas. Thus, schools rarely change their location as long as their facilities do not become unusable. And school authorities are unlikely to build new schools while existing facilities are underutilized. Thus, if a school is closed down and a new one is required to replace the lost capacities, this new school is usually set up in the existing facilities of the closed down school. This is what is done in the CPS. Failing schools can be closed down (CPS Performance Policy). The CPS used this option frequently, closed down schools that were failing or underutilized due to unpopularity, replaced all the staff and often reconstituted schools in the same facilities (WBEZ 2011). Therefore I assume the location of schools to be fixed and that new schools replace closed-down schools in the existing and now free facilities.
6.2.3. Choices regarding institutions

6.2.3.(1) Assignment to home schools

In the model all students are first assigned to a home school. This assignment is based on location, so that each student is assigned to a school that is close to her home. Each student is guaranteed a place at her home school. And any student who does not apply successfully to another school will automatically attend her home school.

This assignment mechanism is a carbon copy of the one used in the CPS. But this approach is also common to all school choice systems that I am aware of. The assignment to home schools has several advantages. It creates an initial distribution that guarantees each student a place close to her home, leaves no school over- or undersubscribed and assigns a school to all students that do not apply to any other school.

6.2.3.(2) Application mechanism

Applications in my model are cost free and all students can apply to any and as many schools as they like. The way I have modelled the application process, utility-maximizing students effectively apply to every school from which they expect a higher utility than from their assigned home school. Then they choose, among the schools that admit them, the one from which they expect the highest utility.

As I intend to analyze the effects of school choice, I have to create an environment in which school choice really happens. For this reason I allow students to apply freely to as many schools as they desire. If individual applications are not costly but may not be successful due to oversubscription, students are likely to apply to many schools. With this approach they can increase their chances to be admitted to at least one school of their preference. In the CPS, students are free to apply to as many schools as they like and applications can be completed without much effort. Consequently, those students that apply at all usually apply to several schools (see 3.3.2.(2)). As distance to schools has a strong impact on applications, the number of schools which any student prefers over her assigned home school is limited. Given the application mechanism in the CPS, it should be possible to fill out applications to all these schools in less than one day.

31 Cost free here means, that there are no fees charged for applying to a school, that the applications forms can be filled out in a short time and that students and parents are not tested or assessed in any way. However, the application to any school usually follows only after information has been gathered about it. This gathering of information is time-consuming and thus costly.
6.2.3.(3) Admission

For students to be faced with real school choice it is not sufficient that they are only allowed to apply freely to schools. It is also necessary that students have a real chance of being admitted at schools of their choice. In my model, admission to schools is independent of student characteristics. Contrary to some existing school choice systems, schools are not allowed to pick or reject students. Neither is the admission mechanism aimed at creating a desired mix of student characteristics at individual schools. Thus, each applicant faces the same probability of being admitted to each school, independent of her own characteristics and of those of the student population at the desired school.

In the CPS, public schools are not allowed to pick or reject students, with few exceptions. As these exceptions are only available to the brightest students and their inclusion would necessitate another layer of complexity in my model, I exclude selective admission schools and the students that attend these schools from my analysis. For the majority of schools that is not allowed to select students, the assignment mechanism in the CPS works as follows: as long as a school has sufficiently many free places to accommodate its applicants, these are all admitted. For oversubscribed schools there is a mechanism that attempts to manage the student mix regarding gender and race. In case of oversubscription, lotteries are held that randomly admit students into the existing places. But not all students enter the same lottery. Students are separated into groups based on grade, gender and race. Then lotteries are held that admit students from each group to the slots that are reserved for this group (Cullen et al. 2003). Each of these lotteries guarantees a random admission to schools. This means that no student is ex ante accepted or rejected with certainty at any school she applies to.

I chose not to replicate this exact lottery system but to hold only one lottery for each school. The main reason for this decision is that it was not possible to obtain information about the number of places reserved for each group. Admission lotteries follow general guidelines but are administered by individual schools. Information about these lotteries is not aggregated. Thus, the staff responsible for data management at the CPS would have had to contact each

32 There are several selective admission public schools that are intended to cater to the most gifted students. In my analysis, these and their students are excluded as are private schools and their students.

33 Some schools also hold separate lotteries for students who live close to the school or whose siblings already attend the school. These lotteries are rarely oversubscribed however. Therefore, I choose to exclude them as Cullen et al. (2003) did. Some schools also have multiple programs which hold a set of lotteries each. As the available student data only identify the attended school, not the attended program, it is not possible to account for such multiple programs.

34 As described above, each student is accepted with certainty at her home school. But when applying to other schools, each applicant has the same chance of being admitted at any school as any other applicant of the same group.
school to gather this information. Moreover, many schools did not use the student ID numbers in the application procedure and used their own forms to gather the necessary data to sort students into groups. Thus, in order to learn the number of applicants in each group at each school, it would have been necessary to either match names or addresses to student IDs or to process the data on the forms issued by the schools. The CPS staff is however not permitted to reveal personal data to outsiders, so all the work on such data would have to be done by CPS staff. Thus, to gather and process the data that is necessary to use individual lotteries for different student types would have taken dozens if not hundreds of work-hours of CPS staff. These would have been associated with prohibitively high costs. Moreover, replicating the original system would necessitate simulating several lotteries for each school. This would further increase the already burdensome amount of necessary computing time.

Moreover, the policy of achieving racially integrated schools seems not to be very effective. Racial integration has improved since school choice was introduced in Chicago in 1988. But in 2006 there are still several schools with above 40% and up to 53% White students and most (98-100%) of the students of about half the schools are either Black or Hispanic (see appendix 2 for where to find figures for individual schools).

The likely consequences on student distribution of holding only one lottery per school as compared to holding several lotteries that are reserved for a gender-race group each depend strongly on the mechanism that assigns places to groups at the individual schools. In the CPS this mechanism is designed to achieve a racially integrated student composition at each school. Specifically, the goal is for each school that enrolls students from outside its assignment area, to achieve an enrolment of 15-35% non-minority (white) and 65-85% minority students (CPS 2008). The overall CPS student population contains about 50% Black, 35% Hispanic, 10% White and 3% Asian students. In theory, the multiple lotteries imply that a student who belongs to a group that is currently underrepresented at an individual school has significantly higher chances of being admitted. But as shown above, the policy is not effective in achieving racially integrated schools. The outcomes are especially devastating concerning the goal of getting non-minority students into schools whose students are predominantly from the minority groups.

This failure to achieve better integration can be due to two reasons. Either the policy is not properly enforced, allowing schools to deviate from the targeted bands. Or there are not sufficiently many applicants of those groups that are underrepresented at most schools. In both cases, the fact that multiple lotteries are held for different groups does not seem to have a strong impact. Therefore, the fact that I only simulate one lottery per school is not likely to
distort the outcomes significantly. Concerning the interpretation of results about sorting by ability it will however be necessary to keep in mind, that students from minorities have better chances of being admitted to non-minority dominated schools should they choose to apply to these.

6.2.3.(4) Programming solution for the admission mechanism

To fully replicate the admission mechanism of the CPS, I would have to store lists for each student and school. The student lists would contain all those schools which the student prefers over her assigned home school. After going through all schools and saving the utility that the student expects for each respective school, I would have to sort each list by expected utility to get a ranking of choice schools for each student. The school lists would hold all applicants to the individual schools. Then a first round of lotteries would be held, one for each school, which admits a random selection of the applicants to the free places. After receiving the outcomes of the lotteries at all schools to which she had applied, each student would pick, among those schools at which she was admitted by the lottery, the school that ranks highest on her list of preferences. Some places would not be taken, because the student who was admitted to this place chose to attend another school. These places would then be assigned by a second lottery or by using waiting lists which hold students that had not won in the first round.

This process would require a strong increase in the already burdensome amount of computing time and memory. Therefore I have decided to take a programming shortcut that delivers the same results but needs considerably less computational resources. The simulation program handles the application and admission procedure as follows:

I compute the expected utility at each school for all students. Then the program cycles in a random order through all students that are about to enter high school. For each student, the program cycles through all schools. For each school that is not already full when it is the turn of the current student, the utility that the student expects from this school is called from memory. If the utility that the student expects from an individual school is higher than

---

35 In the final version, simulating one model run for ten years and two variants of the institutional setting takes about 150 minutes. The 20 runs that are at least necessary to get a usable distribution of outcomes thus necessitate more than 2 days of continuous computing. While programming the simulation, I had to make hundreds of trials with at least 20 runs to check whether the newest piece of code actually does what it is intended to do. The multitude of lists that would be necessary for an exact replication of the assignment process in the CPS (15-17 thousand lists for individual students for each cohort) and the algorithm that finds the currently preferred school and fills all free places through iterations would have increased the necessary computing time, probably at least by an order of magnitude. Even with the three personal computers that were at my disposal I could not have handled such a computational load.
that of the currently favoured school, this school is set as the new favourite school. At the end of the cycle over schools, the currently favoured school is that school which the student likes most among all the schools that still have free places when it is this student’s turn to choose. The student is then admitted to her favourite school and the number of remaining places at this school is reduced by one.

How does this programming duplicate the results of the mechanism used in the CPS? The centrepiece of the programming shortcut is to use the random order of applications by students to represent the random assignment by lotteries that are held at the school level. In the real world lotteries, all those students who preferred the oversubscribed school A and have applied to this school have the same probability of being chosen by the lottery. Therefore all those students who prefer school A have exactly the same probability of being admitted.

In my program, students apply to schools in a random order. Each student who chooses a school is admitted until the school is full. The order in which students choose is random. Whether a student who prefers school A gets a place there, therefore depends only on whether she is allowed to choose before the school is full or not. The order in which students choose is random. Therefore, each student who prefers school A has exactly the same probability of being admitted there as any other student.

Thus, my programming shortcut results in equal probabilities of being admitted at school A for all those students who preferred school A, just as in the case of school-level lotteries. But the computational burden can be handled, contrary to a model that tried to simulate school-level lotteries.

6.2.3.(5) Empty schools
School facilities are expensive to build and maintain. And a school that has no entrants in a year does not use these facilities effectively. Additionally this school has shown that it does not satisfy students and parents. Therefore it is likely not to be continued under the same management. As there is usually not much excess capacity of school facilities in a school district, the school authorities will immediately want to use the facilities that are now free for a new school. Moreover, for programming reasons, it is important to have the same number of schools at the same locations in each turn of a simulation, as in the base year from which the data for this simulated turn are taken. Otherwise, I would have to reassign students, whose original home school is closed in this turn, to other home schools. As I do not have the home
location of students, this reassignment would be hard to do. Moreover, if the attending students left the school, I would have to distribute them to nearby schools, which would make an additional round of applications and admissions necessary.

In my simulation, schools with an empty cohort are therefore reconstituted. This means that the school starts the next year under a new management. And I assume that students who are already attending the school in higher grades stay at the school. A school that attracts no applicant in one year is very unpopular. Such a school was very likely almost empty in the years before. And those students who decided to attend the school in the previous years are very unlikely to choose schools actively and are therefore likely to stay, even if the management changes.

If a school has an empty cohort, it is not possible to generate an official measure for academic quality in the year in which this cohort takes the official test. And for a school that is completely empty, like a newly founded school, it is not possible to compute the other school characteristics, such as a dropout rate or the share of low-income students. But parents and other schools need to have information about newly reconstituted, or newly founded, schools in order to incorporate them into their decision-making process. I assume that parents and other schools replace the non-observable characteristics of newly founded or reconstituted schools with the district-wide averages. That is, they expect a new or reconstituted school to be of average quality.
7. Data and Calibration

7.1. The Chicago Public Schools

The Chicago Public Schools (CPS) is one large school district that covers the inner city of Chicago. This district contains dozens of high-schools and thousands of students in each cohort in a relatively small area. As a consequence of this high density of schools, there are several schools within commuting distance of each student home. In the CPS, each student can apply to all public schools, and most public schools assign applicants to free places by lotteries. Financing follows students, so that schools that loose students also loose funding and have to reduce staff. Moreover, schools that are underutilized can be, and actually are, closed down and reconstituted under a new management. This setting of the school choice system is favourable to a strong impact of school choice. And this setting of free school choice has been in place in a similar shape for many years, so that schools and parents had time to adapt fully to its conditions.

A wealth of data is centrally collected in the CPS. Students receive an identification number, by which they can be tracked throughout their schooling career. This makes it possible to match information on attended schools to individual characteristics and results in centrally administered tests.

School data are also collected and made publicly available in ways that are easily accessible to parents. The official measure for the academic quality of schools is published in the form of rankings in newspapers and on the internet. School report cards summarize characteristics of the schools, such as the official quality measure, dropout rates, the share of English language learners or the racial composition of schools. Moreover, the homepage of the CPS provides a wealth of information on school characteristics for parents who are willing to search for information beyond what is summarized on the school report cards. With the high number of schools and programs available to students at the high-school level, it is likely that parents base their school choice decisions at least partially on official information of school characteristics, especially if such information is easily accessible.

Previous literature has shown that parents do use official information on school characteristics, if such information is available. I will confirm this finding for the CPS in 8.3. What is important to note here is that, using historical data on these school characteristics, I
have access to the very same information on many school characteristics that parents had at their disposal when they made their school choice decisions.

7.2. The data

7.2.1. Description of the data

7.2.1.(1) The student data

A dataset containing student-level data was provided upon request by the Office of Research, Evaluation, and Accountability (REA) of the CPS. This dataset covers all students who entered the 9th grade of the CPS in the cohorts 2001-2006, who were still enrolled at the end of the 9th grade and were supposed to start the following school-year in the CPS. This means that students who left the CPS during their first year at high-school - because they moved away, went to private schools or dropped out - are not included in the dataset.

The dataset contains a student identifier. By using this identifier, students can be tracked throughout their school careers. The dataset also contains individual student test results in two centrally administered tests. The first test is the Illinois Standards Achievement Test (ISAT) in grade 8 for math and reading. I can therefore use its results to approximate the innate ability\(^{36}\) (innate) of students at high-school entrance. To compute the variable for innate ability, I use the average of the test results for math and reading for each student. Based on this variable, I assign students to ability levels. High-ability students (high_innate) are the 20% of students with the highest scores in the ISAT, low-ability students (low_innate) are the 20% of students with the lowest scores in the ISAT. Both groups are identified by dummy variables. Medium-ability student are the reference group that contains the remaining 60% of students. I also aggregate the average ISAT test result of individual students by school and cohort and use this number to measure average innate peer ability (innate_peer).

The second test for which results are included in the dataset is the Prairie State Achievement Examination (PSAE) in grade 11 for math, reading and science. This test is taken in the third grade of high-school. It is thus affected by the quality of education of the attended school and I use the results in this test to measure the individual academic achievement of students.

\(^{36}\) This innate ability represents the academic ability of a student at the time of high-school entry as measured by centralized tests. This innate ability includes the effect of previous schooling. And it is a narrow definition of the students' ability. But the student achievement that I intend to explain and to predict using this innate ability is the individual outcome in another centralized test. For this intention, the result in a previous test is a suitable proxy.
To compute the measure for outcome for an individual student, I use the average of the three test results for math, science and reading of this student. Below I will compute the quality measures for schools based on these individual student outcomes.

The dataset also contains information about the status of the student in the years after high-school entry. There are variables that cover whether a student was still active in the CPS in the second and third year after high-school entrance (status_2ndyr, status_3rdyr). For inactive students there also is a leave code that identifies whether each student who left the CPS dropped out, left the area or left to a school that is in the area but outside the CPS, like a private school (lv_code2ndyr, lv_code3rdyr). Another set of variables covers the attended grade in the second and third year after high-school entrance (grade_level_2ndyr, grade_level_3rdyr), which makes it possible to see whether a student repeated grades.

The dataset also contains a set of individual student characteristics, which cover race (race), gender (gender), special education status (iep), whether the student was eligible to receive lunch for free or at reduced prices (lowincome) and whether the student had English language learner status (ell). Note, that the dataset only includes students whose English was deemed good enough to participate in the ISAT exam. Their peers whose English was worse took another test whose outcomes were not included in the sample. The birth year is also included which I use to compute the age at high-school entry. I use the age at entrance as a proxy for grade repetition in elementary school (rep_elem).

Another set of variables in the student-level dataset holds information that is crucial to the analysis of choice behaviour. These variables contain the assigned schools for grades 8 and 9 (aunit_grd08, aunit_grd09), the school where the student took the ISAT (isat_unit_tested) and the attended schools in grade 9 through 11 (unit_number, unit_2ndyr, unit_3rdyr). The assigned school is the home school which the student attends by default, if she does not apply successfully to another school. The information on the assigned school allows two insights. Firstly, as the home school is always close to the student’s home, I can use the location of the assigned school as a proxy for the home location of the student. Secondly, the assigned school allows me to identify which students did actively choose schools. If the attended school in grade 8 is not identical to the assigned school in grade 8, I know that the student has actively chosen her elementary or middle school (elem_changer). If the attended school in grade 9 is not identical to the assigned school in grade 9, I know that the student has actively chosen her high-school (hs_changer). The identification of the attended and assigned schools also allows me to link the school datasets to the student dataset. Moreover, this identification of assigned
and attended schools allows me to aggregate several student-level variables at the school level, which are not included in the school-level datasets. Most important of these variables for the simulation and the preparatory empirical work are the number of assigned and attending students for each school and cohort and the average innate peer ability.

The type of the attended school is also included in the student-level dataset. Based on this school type, I can identify whether a school was allowed to use selective enrolment or whether it had to use indiscriminate lotteries in order to assign applicants to free places.

7.2.1.(2) The school data

I use three datasets on the school level, which were all freely available on the CPS homepage (CPS School Data)\(^{37}\). The first dataset contains average school characteristics for the school years 1989 through 2006 for all schools in the CPS. This dataset contains the following variables that are aggregated at the school level: Dropout rate, graduation rate, attendance rate, truancy rate, mobility rate, the share of students of each race (African American, White, Hispanic, Native American, Asian/Pacific Islander), the share of students who receive lunch for free or at reduced prices and the share of students who are English language learners.

A second dataset contains the official quality measure for the years 2001 through 2007. This quality measure is the share of students at a school that met or exceeded the state standards in the PSAE test in grade 11.

A third dataset contains the probation status of schools for the years 1997 to 2010. This probation status is based on performance of the schools in the previous years.

In addition to the official datasets, I needed to identify the location of the schools. This information was not provided by the CPS. I had, however, the full name of the schools. Based on these names, I identified the postal addresses of the schools. In a second step, I identified the longitude and latitude of these addresses. Based on the geographic location, I computed the matrix of distances between all schools that were included in the final dataset\(^{38}\).

\(^{37}\) The data on this homepage has been updated meanwhile and the included datasets have changed. I therefore include the original and unchanged datasets on a CD-ROM that accompanies the dissertation.

\(^{38}\) This matrix contains air-line distances between schools. These are of course only approximations of real distances and travel-times between schools. As there are more than 60 schools in my sample, the total number of distances between any two schools, which is \((n-1)+(n-2)+(n-3)+\ldots+1\), is close to 2000. The effort to compute exact travelling distances or travelling times, for example by using the “Get directions” tool of Google Maps would have been prohibitively large and I do not know how to employ a geodata-based computation of distance matrices. Moreover, as most students do not attend schools that are very far away and due to the checkerboard pattern of most streets in Chicago, air-line distances are a reasonable approximation of travelling-distances.
In order to identify the share of places that remain free at a school after the choice process is completed, I need some information about the number of students that a school can accommodate for one cohort. Information on this capacity was not provided by the CPS. What I had information about is the number of entrants in each year and the number of assigned students for each year. I assumed, that no school ever took in more students than it could accommodate, and that the CPS never assigned more students to a school than this school could accommodate. Based on these assumptions, I assumed the capacity of each school to be the maximum of the following two numbers: (a) the highest number of students that were taken in at this school in any one year that was covered in the student-level dataset (b) the highest number of students that were assigned to this school in any one year that I can observe.

7.2.1.(3) Combining the datasets

I retrieved all school-level data from the various datasets and stored them in matrices, so that they can be accessed based on school-id and year. Then I combined the datasets, so that all information is included in a student-level dataset, either in the dataset directly or in matrices that are stored with the dataset. Most school characteristics are included in the form of additional variables in the student-level dataset, both for the cohort that the student attended and for previous cohorts and both for the assigned and for the attended schools. I will later use these variables of school characteristics when I estimate the education production function and the decision functions of the actors.

The entire preparation of the dataset, including the cleaning of the data that is described below, is done with a set of do-files in STATA which starts with the original datasets as they were provided by the CPS. This approach makes it easy to change the process of data preparation, if necessary. Moreover, this approach makes it easy to recapitulate the process of data preparation.

7.2.2. Cleaning the data

The original datasets contain observations which are either obviously false, or which I cannot include in the estimations or in the simulation for various reasons. In this sub-section I explain which observations I have excluded and why I excluded them. Based on the name of the

---

39 For some of the datasets, a conversion into a format which can be processed by STATA was necessary. As such a conversion cannot be done in STATA, as far as I know, this conversion is not documented in the do-files. I used StatTransfer for all conversions. The original files are also included on the accompanying CD-ROM in the original data format.
STATA command “drop”, I use “dropping” an observation as a synonym for excluding an observation.

7.2.2.(1) Student data
I excluded all student observations whose school type or attended grade does not fit the dataset. These dropped observations include students attending an “early childhood” school and students who seem to have attended a grade lower than 8 or higher than 12 in the second or third year after entering high-school. I also dropped students whose high-school entering age was smaller or equal than 13 or higher or equal than 19. These observations either contain data glitches (high-school freshmen at an age below 10 or above 40 were included) or identify students who are exceptional in entering high-school very early or very late, so that I do not want these observed values to distort the effect found for students who entered high-school at, or close to, the intended age.

I had to drop another group of students because the crucial information on assigned schools is missing. If I cannot observe the assigned school for grade 9, I do not know whether a student has actively chosen her high-school.

I also dropped all Native American students. There are only 106 Native Americans students in a sample of more than 80,000 students who are left after exclusions for all other reasons. I will later differentiate several estimations by race and cohort. If I wanted to include Native Americans, these estimations would have to be done based on less than 20 students for each cohort, which is too low to get reliable estimates.

Finally, I dropped students who were assigned to schools that are no home schools. As these schools are no general home schools, the assignment of only a few students has to be based on student characteristics. Schools with a very low number of assigned students are mostly charter schools with less than 20 assignees in one cohort. Students who are assigned to small charter schools are most likely the children of the founders of the charter school. These students are likely to have parents who are much more involved in the education of their children than the average parent. And these students have a guaranteed place that is not open to all other students. Any other assignments to schools that are no home schools have to be based on intervention, most likely of the parents. About the reasons for such exceptions I can only speculate. The most likely reasons are certain disabilities, that prevent travelling far or that can be accommodated only at schools with specialized programs. All the students who are assigned to non-home schools are therefore likely to differ systematically from their peers and/or are unlikely to choose another school.
7.2.2.(2) School data

I excluded all selective enrolment schools. In the CPS these are College Prep Schools, Selective Enrolment High-Schools and Military Academies. These schools all select their students based on the characteristics of the applicants. The exclusion of these schools from my simulation has several reasons. The first reason is, that I cannot replicate the selection procedures of the selective enrolment schools properly. This selection process is based on 7th grade scores in centralized tests, on 7th grade classroom grades and on the results in an entrance exam (CPS Selective Enrollment). I have access to none of these data and could only approximate the 7th grade test scores with an acceptable precision. Moreover, there is the possibility to be admitted via principal’s choice, a process for which I cannot even approximate the criteria. The second reason for the exclusion of selective enrolment schools is that a simulation of school choice is already complicated and computationally burdensome if only the choices of the students are simulated. If I were to simulate selection of students by schools as well, the computational burden would increase even further. The third reason is that selective enrolment schools cater to a different clientele than freely accessible public schools. Students who get into selective enrolment high-schools differ systematically from their peers. They have to have better grades, have to be willing to take an extra admission exam and have to be willing to face the competitive environment in such schools. Moreover, these students have an outside option that most of their peers do not have.

I therefore excluded selective enrolment high-schools from my analysis. And I excluded the students who attended these schools. As presented above, these students are a clientele that differs systematically from the students who attend freely accessible public schools. Therefore, as long as I apply my findings only to those students who stay in non-selective public schools, it should be possible to exclude those students who attended selective enrolment schools without distorting the results. Moreover, including students from selective-enrolment schools while excluding their schools would distort the simulation in two ways. A higher number of students would have to be fit into the capacities of the non-selective schools. And the influx of academically strong students would alter average student characteristics at the remaining schools.

On the opposite end of the distribution of innate academic ability, I also excluded Academic Prep Centres. These schools catered exclusively to students who were too old to remain in elementary school, but academically not fit to enter a high-school. Most students left these schools within one year, either to drop out or to be transferred to a normal high-school. Moreover, this type of school was discontinued after the cohort of 2002 and barely exists in
the period which I analyze. Following the same reasons as for selective enrolment schools, I excluded Academic Prep Centres and the students who attended them.

The third type of schools that I excluded is schools that are very small. I exclude all schools with a total of less than 200 entrants in the 6 cohorts of my dataset and all schools with a total of less than 400 entrants in the 4 cohorts for which I have full data (2002 through 2005). But I did not exclude schools with a total of less than 400 entrants (in 2002 through 2005) if these were regular home schools.

There are several reasons for the exclusion of small schools. The first reason applies only to the exclusion of schools with less than 200 entrants. By their unit-identification numbers I can identify all these schools as charter-schools. Most of these tiny charter schools are newly founded. Founding a charter school is a lot of work and is usually only done, if the founders see a need for a school in the area that caters to a special group of students. Small charter schools are therefore usually focused on special needs or interests and are not attractive to the majority of students. Moreover, the children of founders usually have a guaranteed place at a charter school, so that the share of places that is free to outside applications is rather small. As small charter schools cater to specific student types, their attending students are likely to be systematically different from the average student in the CPS. Moreover, these students have an additional choice option and are unlikely to apply to any other school. I therefore excluded both these tiny schools and the attending students.

There are two reasons for the exclusion of schools with a total of less than 400 entrants in the 4 cohorts of my dataset. The first reason is the computational load. For each school that is active in the simulation, I have to compute the attractiveness of this school to each student, the competitive threat this school poses to all other schools, the average school characteristics and the quality measure. Moreover, I have to predict a productivity level and I have to repeatedly store and retrieve all these values from matrices. A doubling of the number of schools would therefore at least double the necessary computing time. This increase in the computational load is not worth the small increase in the number of places for entrants in a cohort that would result from retaining these small schools. The second reason for excluding small schools is system stability. Parental choice depends on school characteristics and the productivity choice of schools depends on the competitive threat from other schools. The smaller a school is, the more its characteristics can be affected by the characteristics of a few students. One exceptionally bright student could have, for example, a strong impact on average test scores of a class of 15 students, while her test results would hardly affect the average of a high-school where she has 300 peers. Small schools therefore show a higher
variation in student characteristics and create a lot of noise in the decisions of parents and schools.

I therefore excluded schools with a total of less than 400 entrants in the 4 cohorts for which I have full data unless they are regular home schools, to which all students in the surrounding area are assigned.

Compared to the tiny and very small charter schools, these small home schools are small, but fairly normal. The characteristics of their students also do not differ strongly from their peers at bigger schools. I therefore excluded neither these small home schools nor the students who attend them.

### 7.2.3. The final dataset

After dropping students and excluding types of schools for various reasons, I have a final dataset of 82346 students and 62 high-schools. This dataset contains the students of the cohort of 2006 (and a few students from the cohort of 2001), who are only included in some preparatory estimations. The dataset of students in the four cohorts for which I have full data, and which will populate the simulations, contains the same 62 schools and 65665 students with a cohort strength between 15759 and 17076. As can be seen from the tables below, the differences in average characteristics between the two populations are minimal, as should be expected because the only relevant difference between the datasets is that an entire cohort (of 2006) is included in the first dataset but not in the second.

Students who enter preparatory estimations:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>82346</td>
<td>.5086586</td>
<td>.4999281</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lunch_status</td>
<td>82346</td>
<td>.8783912</td>
<td>.3268355</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>iep</td>
<td>82346</td>
<td>.1755641</td>
<td>.3804512</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ell</td>
<td>82346</td>
<td>.0187744</td>
<td>.1357283</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_WHITE</td>
<td>82346</td>
<td>.0747213</td>
<td>.2629427</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_African_Am</td>
<td>82346</td>
<td>.5416899</td>
<td>.4982619</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_Asian_Pac</td>
<td>82346</td>
<td>.0250164</td>
<td>.1561758</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_Hispanic</td>
<td>82346</td>
<td>.3585724</td>
<td>.4795842</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_changer</td>
<td>82111</td>
<td>.3307109</td>
<td>.470472</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>hs_changer</td>
<td>82346</td>
<td>.4786996</td>
<td>.4995491</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>innate</td>
<td>82346</td>
<td>152.6666</td>
<td>11.16696</td>
<td>120</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 7-1: Average characteristics of students in the dataset for preparatory estimations

---

130
Students who enter the simulation:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>65665</td>
<td>.510683</td>
<td>.4998897</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lunch_status</td>
<td>65665</td>
<td>.8775908</td>
<td>.3277603</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>iep</td>
<td>65665</td>
<td>.1756034</td>
<td>.3804853</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ell</td>
<td>65665</td>
<td>.0198431</td>
<td>.1394621</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_White</td>
<td>65665</td>
<td>.0763573</td>
<td>.2655709</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_African_Am</td>
<td>65665</td>
<td>.5425874</td>
<td>.4981868</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_Asian_Pac</td>
<td>65665</td>
<td>.0255692</td>
<td>.1578473</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>r_Hispanic</td>
<td>65665</td>
<td>.3554862</td>
<td>.478664</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_changer</td>
<td>65665</td>
<td>.337303</td>
<td>.4727929</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>hs_changer</td>
<td>65665</td>
<td>.4915861</td>
<td>.499933</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>innate</td>
<td>65665</td>
<td>152.4339</td>
<td>11.15811</td>
<td>120</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 7-2: Average characteristics of students in the dataset for the simulations

As can be seen from table 7-2, 51% of the students who enter the simulation are female, 87.8% receive free or subsidized lunch (lunch_status=1), 17.6% receive special education. More than half the students (54.3%) are African American, about a third are Hispanic (35.5%), only 7.6% are White and 2.6% are Asian or Pacific Islander. From data on the attended and the assigned schools I can see, that about one in three students (33.7%) actively chose her elementary school (elem_changer and about every second student (49.2%) actively chose her high-school (hs_changer). Only 2% of the students are English Language Learners (ell). Note that only those ell students are included whose English was good enough to take the ISAT.

These statistics do not represent an average US student. About 90% of the students are from “minorities”, almost as high a share receives free or subsidized lunch and close to a fifth receives special education. These characteristics describe a group of students who are not very likely to do well in school. But it is a group of students which is to be expected in a school district in the inner city of a metropolitan area in which there exist extensive outside options for those students whose grades are high enough or whose parents can afford to pay tuition at private schools. The students in my final dataset are those who have no outside options. These students are trapped in the public school system and are those, who are most strongly affected by the conditions that apply to the public schools. I will have to keep the specifics of this group of students in mind when interpreting findings in estimations and simulation outcomes.
8. Preparatory empirical work

In this chapter I will calibrate the simulation of a school choice system. This calibration will be based on the dataset that resulted from the preparations in chapter 7. The central pieces of this calibration are the estimation of the decision functions for all actors (students/parents and schools) as well as the education production function (EPF). Once the simulation is calibrated, I can utilize it to analyze the impact of variations in institutional settings and/or conditions of the school choice system.

8.1. Educational productivity of schools

I have defined the educational productivity of a school as the effect of the quality of education on academic outcomes at this school. This quality is measured as the academic achievement of students at the school, as compared to the district-wide average achievement of students who show comparable characteristics and who faced a comparable influence from peers. This rather abstract concept can be measured as the average deviation of students at a school from the scores that they were predicted to score. If students at a school score on average better than their peers who face similar conditions concerning student characteristics and peer characteristics, a school has a high educational productivity.

In this subchapter I will compute the educational productivity of all schools which are included in my final dataset for each cohort of students. For this computation I will use an approach similar to the approaches used by Cullen and Jacob (2007) and Lauen (2007a) for the CPS, described by Wilson and Piebalga (2008) for schools in the UK and by Meyer (1997).

The approach to computing the educational productivity works as follows. I will first estimate an auxiliary education production function for the student-level dataset that explains the outcome in the 11th grade PSAE test with 8th grade ISAT test results, student characteristics and the average innate ability of students at the school (8.1.1.). Based on this estimation, I will predict test outcomes for all those students for whom I can observe real world test results (8.1.2.). Then I will compare the predicted outcomes to the real world outcomes. I define the average of this deviation of real world outcomes from predicted outcomes at a school as the educational productivity of this school (8.1.3.).
8.1.1. Estimation of the auxiliary education production function

The first step in the computation of educational school productivity is to estimate the auxiliary education production function. For this estimation, I have to choose the set of variables which I use to explain the outcome, the 11th grade PSAE test results.

The aim of the simulation in chapter 9 is to identify the possible real world effects of using an alternative measure for school quality which is based on school productivity. My aim in choosing the set of variables is therefore to identify a set that might be used in reality to compute an official measure for school productivity. Previous literature that described or used value-added measures can be used as a starting point for the identification of such a set of variables.

The value-added measure that was introduced in the UK in 2002 accounted only for prior academic achievement. A newer version that was introduced in 2006 additionally accounts for the following student characteristics: gender, special education status, whether the student received free lunch, ethnicity, foreign mother tongue, mobility, within-year age, whether the student was “in care” while at school and a measure for the deprivation of the student’s home neighbourhood (Wilson/Piebalga 2008). Meyer (1997) proposes to estimate school productivity by using slope parameters for the impact of: prior achievement, individual student characteristics and school-level factors. Cullen and Jacob (2007) calculate a value-added measure for schools in the CPS by accounting for prior academic achievement and the following student characteristics: race, gender, age and eligibility for free or reduced lunch. Lauen (2007a) uses a value-added measure that controls for prior achievement, fixed effects for the quality of the prior school and student background statistics including the eligibility for free or reduced lunch, gender, special education status, within grade age, the number of prior school moves and a proxy for the distance to the neighbourhood school.

In the estimation of the auxiliary EPF I include innate ability, approximated by the 8th grade ISAT test results. I also include the following student characteristics: gender (gender), special education status (iep), English language learner status (ell), whether the student receives lunch subsidies (lunch_status), and dummy variables to identify race (drace_afram, drace_asiapac, drace_hisp). These variables include most of the information that I have on individual students. Of the list of variables used in the UK (Wilson/Piebalga 2008) and/or by Lauen (2007a), my list only misses those variables for which I do not have student-level data. The variables used by Cullen and Jacob (2007) are all included.
Student characteristics that I can observe but that I do not include are whether the student has actively chosen elementary school (elem_changer) or high-school (hs_changer). Active choice could be used as an indicator for how important education is to the parents. But the original dataset does not contain the assigned elementary school for the entire cohort of 2002. Therefore, I would loose this cohort for the estimation if I included elem_changer. More importantly, the changer-variables were not used by previous literature that tried to compute school productivity as these variables necessitate information that is not readily available in many school choice systems and are therefore unlikely to be used in reality. Moreover, including these variables only leads to a slight increase in the R-squared.

Additionally to innate ability and the individual student characteristics described above, I include the average innate ability of the peers at the attended school (innate_peer). The ability of peers might have an impact on individual student outcomes. If there is such a peer effect and I do not include a variable on peer ability to capture at least some of its impact, this omission might distort the measure for productivity that I intend to capture. If a high peer ability increases student outcomes, then this impact results in a positive effect on scores at a school with a high average innate ability. If I do not include peer ability in the estimations, this positive impact on scores would remain in the unexplained residual and would thus be attributed to the unobservable educational productivity.

Due to the inclusion of peer effects in this simple form, by including only innate_peer in an OLS-estimation, the effects of peer ability in my simulations are linear and independent of other student characteristics. I recognize that this representation of peer effects is rather simplistic, but I chose this approach for two reasons: First, there is an extensive literature on the strength and shape of peer effects, that I cannot cover satisfyingly as yet another field of the economics of education within the limitations of this dissertation. In order to model the peer effect in a more refined way, I would have to choose a shape based on one theory in the literature or on the findings of one study, without the capability to judge whether it is the most realistic one. Alternatively, I would have to choose several shapes for the impact of peer effects and run them in alternative simulations, thus multiplying the already burdensome computing time. The second reason not to use a more sophisticated representation of peer effects is, that I intend to use a value-added measure that is likely to be used in reality by school authorities. As the concepts of value-added and of predictions based on regression results are already not easy to grasp for parents, teachers and principals, I prefer not to deviate from the simple OLS-approach for estimating the EPF.
I did not include average school characteristics like the dropout rate, truancy rate and lowincome rate in the estimations for both the auxiliary EPF and the real EPF. This exclusion is due to a few reasons that apply to individual variables and mostly to a general concern about these school characteristics. Concerning the truancy rate, I do not have student level observations on truancy. Therefore, it is not possible to determine the truancy rate at a school, once the simulated school choice has resulted in a different distribution of students to schools. Thus, I could not base predictions of outcomes in the simulation on an estimation that includes the truancy rate. Concerning the dropout rate, I found preliminary evidence that some schools might try to increase the dropout rate in order to keep weak students out of the centralized tests and thus increase average test outcomes (see below in 8.9). If that were the case, the dropout rate would reflect not only average student characteristics but also school policy that differs across schools. Special education status (IEP) covers a wide range of student types. This range includes for example blind or deaf students, students with learning disorders and students in wheelchairs. Many schools are specialized in catering to certain kinds of IEP students, so that the different types are distributed unevenly. It is not unlikely that a student who is blind or in a wheelchair scores as well in school as her non-IEP peers. Other IEP student types, for example those with learning disorders, are however likely to score worse than non-IEP students. Moreover, some IEP students are taught in separate classes, as would be necessary for example for students who are simultaneously deaf and blind. Others share the classroom with non-IEP students and have a much stronger impact on these peers. Without knowledge about the distribution of types of IEP students and the extent of their inclusion in classrooms with non-IEP students, the simple share of IEP students at a school might be a misleading information about the learning environment at a school. The most important reason for not including school-level averages of student characteristics such as truancy rate, dropout rate, lowincome rate and IEP rate is however autocorrelation. Student who have a low income, an IEP status, are truant or more likely to drop out have on average lower test scores in grade 8 and thus a lower innate ability. Thus, there would be a high correlation between these aggregate measures and average innate ability, leading to autocorrelation problems in estimations if these variables were included.

I will later measure educational school productivity as the average difference between predicted test results and real world test results at the school-level. This average deviation can be affected by overall school quality and by classroom effects. One example for classroom effects is the teacher. An exceptionally good math teacher could result in a positive difference between prediction and real world outcomes for math, even at a school which has a low
overall educational productivity. As it is not rare for a teacher to be teaching, for example, “10th grade math” it is likely that many or even most students at one school were instructed in math by an individual teacher for some time during grades 9 through 11. Given the data, I cannot separate school- and classroom-effects. But I can reduce the impact of classroom effects by using the average deviation for the three outcomes for math, reading and science. Students were in different classes for the three measured outcomes. Therefore, classroom effects in one subject can affect only a third of the measure for educational school productivity and it is harder for classroom effects like the exemplary brilliant 10th grade math teacher, to distort the measures for overall school quality.

The predictions for math and reading use the corresponding 8th grade ISAT results for innate ability. As there is no science outcome reported for the ISAT, I also use the math results of the ISAT as innate ability for the explanation and prediction of the science outcomes. In order to compute the productivity for each school and each cohort, I estimated the auxiliary EPF and predicted outcomes separately for each of the three subjects and each of the 5 cohorts that were covered by my data. I cannot present all 15 estimation outcomes here. However, the insights that can be gained from these estimations are interesting. Therefore, I will present one estimation table for the auxiliary EPF of the aggregate outcome that covers all cohorts and all groups of students. Additionally, I will include a table that presents the variable estimates for the three subjects science, math and reading separately.40

---

40 The interested reader can easily access the estimation tables by subject by running the do-file “Table for 8.1.1. estimation EPF to get productivity.do” in the folder C:\Data\Stata CPS Data\Graphs and Tables
As can be seen from the estimation results in table 8-1, innate ability is the most important driver of academic outcomes. An increase of 1 point in 8th grade test scores leads to an increase of 0.78 points in 11th grade test scores. The standard deviation for the aggregate outcome is 11.39, which I will use to put the estimates for the other variable estimates into perspective. Peer ability also has a strong impact. An increase of 1 point in average 8th grade test scores of the peers leads to an increase of 0.15 points in 11th grade test scores. Girls score slightly worse than boys. At first glance this result conflicts with common findings, but is due to the fact that the composite outcome includes two subjects where boys usually score better, science and math, and only one subject where girls usually score better, reading. Students who receive free or subsidized lunch (lunch-status) score 9% of a SD worse than their peers. Students who have a special education status (iep) score 21% of a SD worse and English Language Learner status (ell) does not have a significant effect. Compared to the reference group, the White students, African American students score 20% of a SD worse, Asians and Pacific Islanders score 8% better and Hispanic students score 11% of a SD worse.
Table 8-2 shows the variable estimates for estimations by subject. Here I will only comment on differences between the estimates. Based on these results, it is obvious that the uncommon finding that girls score slightly worse than boys is an aggregate of higher scores for boys in science and math and higher scores for girls in reading. The negative impact of lunch_status is weaker for math than for reading and science. Similarly, the disadvantage of African Americans and Hispanics is least pronounced for mathematics while the advantage of students of Asian and Pacific Islander students is only significant for mathematics. The impact of the average innate ability of peers (peer_math, peer_read) is much weaker for mathematics than it is for science and reading. Finally, the impact of own innate ability (innate_math, innate_read) is stronger for reading than for math. The impact for science has to be treated cautiously, as I do not have data on test results for 8th grade science. Instead I used 8th grade math test results. The use of a proxy is likely to decrease the impact of innate ability for science.

8.1.2. Prediction of test results

Based on the estimation results for the education production function in 8.1.1., I predict test scores for all students for whom the dataset contains real world test results. I have to limit predictions to those students for two reasons. The first reason is, that for students without real world test results, I cannot compute the individual deviation between prediction and real world outcome. Predictions for these students would therefore be useless in computing the average deviation at a school. Theoretically, I could compute predicted test for all students,
aggregate these at the school level and compare the average predicted results to average real world test results for the same school. But this comparison would be distorted due to the second reason; the students who miss real world test results in the dataset do not miss these results by pure random chance. Large shares of those students who miss real world test results are dropouts. Students who dropped out in reality have, on average, characteristics that are unfavourable to high scores in tests. Therefore, predicted scores for these students would be low. Including this group of weak students only in the computation of aggregate outcomes for the prediction case would result in an average negative deviation for most, if not all, schools.

In order to predict test results for those students for whom the dataset includes real world test results, I use the STATA-command “predict”.

This command computes a linear prediction for the test result, based on the estimated coefficients of the auxiliary EPF and on the values of the corresponding variables for the observation. This means, that predictions are fully determined by variable values. Consequently, there is no error term in the predictions that are generated in the simulation for the probability to apply to any school, for the function that determines the expected utility from an individual school, for the auxiliary and final EPF and for the educational productivity function. Using the command “predict” has, compared to the alternative command “uvis” an advantage that is central to my approach to achieve robustness for the simulations and a disadvantage concerning the EPF. I will present the alternative command “uvis” and its advantage and disadvantage in comparison to “predict” in a little more detail than would be necessary here. The reason is, that the same reasoning will apply again to the use of “predict” or “uvis” in the prediction of dropout in 8.9.

The disadvantage of “predict” results from the linear nature of this command and from the lack of any random element in the predicted values. A student with a given set of values for the explanatory variables is always assigned the same predicted outcome. Therefore, predicted scores for students with academically weak characteristics are always low and predicted scores for students with strong characteristics are always high. The official quality measure for schools in the CPS is the share of students who meet or exceed state standards. The fact that this official quality measure divides students sharply into those that score above and below a single threshold creates difficulties when combined with a prediction that has no random element. The use of the linear prediction “predict” amplifies the differences in this quality measure between schools with an academically weak and an academically strong student clientele. The mechanism behind this amplification works as follows:
In reality, the test results of students deviate from the average value that can be expected given their characteristics (=the prediction). In a school with mostly weak students, some of the weak students score in reality worse than expected. This has no effect on the quality measure, as both the expected and the real outcome are below state standards. But some weak students score better than expected and do meet the state standards against expectations. Of those students who are expected to score above standards some score exceptionally well. Again, this is not reflected in the quality measure, as both the real and the expected result are above the state standards. But some of the strong students will score below expectations and lower the official quality measure. In a school where the majority of students is expected to score below standards, the random variation results in a net effect of a few more students crossing the threshold upwards than downwards, as compared to predicted outcomes, and thus increases the measured quality of the school. With predictions based on “predict” this is impossible, and the quality measure for the school will be a little worse than it is in reality. For schools with a majority of strong students, the pattern is reversed. Thus, when using “predict” and based on the share of students who meet state standards, schools with a majority of strong students appear a little stronger and schools with a majority of weak students appear a little weaker than they are in reality. Hence, the utilization of “predict” amplifies differences between schools with academically weak and academically strong students.

Predictions using the command “uvis” include the prediction of an error term. This predicted error term has a distribution that is intended to replicate the distribution of the error term in the original data. When using “uvis”, some weak students therefore do meet state standards and some strong students do fail to meet these standards. Thus, “uvis” avoids the amplification described above. But avoiding the amplification would come at a great cost regarding the robustness of the results of my simulation.

The cornerstone of my robustness approach is the random variation of the estimated coefficients of the EPF in each run of the simulation. This random variation is only possible for the command “predict” as this command creates estimation results that can be stored, read and, most importantly, altered. The alteration of simulation results could be used to falsify empirical findings. Therefore, the program STATA includes security measures that are meant to prevent any such alteration. These security measures can be sidestepped, using a programming trick. This programming trick only works for “predict” and not for “uvis”, as

---

41 By changing the stored results of an estimation, it would be possible to falsify estimation results. Therefore, the stored estimation results are given a high security level in STATA, so that they cannot be altered manually. It is possible to write a small command and assign to this command the same high security level. Then, this
the underlying estimation results for “uvis” are not stored in an accessible way. Faced with a choice between the loss of the cornerstone of my robustness measure concerning the EPF and a slight amplification of the effect of an exceptionally strong or weak student clientele, I choose the latter.

8.1.3. Computation of the educational productivity and comparison to average outcomes

The variable for the educational productivity of schools is computed by taking the average difference between predicted and observed test results. I aggregate these differences by school and the year in which the PSAE test was taken. This approach results in a matrix that holds the average deviation of students at each school in each test-taking cohort. Values above 0 mean that students at this school and in this cohort scored on average better than their peers who showed similar individual characteristics and faced similar average peer ability.

The graph below (Figure 8-1) shows the distribution of the variable for educational productivity. Note that this distribution is based on the students and shows the educational productivity of the attended school for each student. Thus, if school A had 300 students in cohort X with an educational productivity of 0.535, and 250 students in cohort Y with an educational productivity of 0.602, it enters into the distribution that underlies this graph as 300 observations with a value of 0.535 and 250 observations with a value of 0.602. Many schools accommodate hundreds of students in each cohort, who then enter this distribution with hundreds of observations with exactly the same value. One result of these multiple identical values is the saw-tooth nature of the histogram.
The distribution of educational productivity has, by definition, a mean very close to 0 at -0.004 with the bulk of the observations between -2 and +2 and shows a shape that is close to a standard normal distribution\textsuperscript{42}. The standard deviation is 1.19 and the spread -3.32 to 3.98. This means that students at the best school scored on average 35% of a SD higher than predicted, based on their individual characteristics and average peer ability. Students at the worst school score 29% of a SD worse than predicted. One SD of the measure for educational productivity is equivalent to the students at this school scoring 10.4% of a SD of outcomes above the predicted value.

It is possible that the measure for educational productivity contains the effect of unobserved student characteristics. If education is seen as important at the home of the student, for example, this is likely to have a beneficial effect on test scores. If students with an unobserved characteristic that is beneficial to test scores accumulate at a school, this will show up as a higher educational productivity for this school. Such an accumulation would be especially worrisome, if it were a reaction to a high educational productivity: If productivity were

\textsuperscript{42} The do-file “Table for 8.1.3. distribution of productivity.do” produces a “pnorm”-plot which shows that educational productivity is normally distributed, a “qnorm”-plot that shows that educational productivity is standard-normally distributed within the spread of observed values and the output of a command that computes skewness and kurtosis for the distribution of educational productivity (0.11 and 3.34) which are close to the values for a standard normal distribution (0 and 3).
observable, it would be likely that those students whose parents see education as important apply more often to highly productive schools. Consequently, schools which at some point in time showed a high educational productivity would systematically accumulate students with unobservable characteristics that are favourable to test scores. This accumulation would then artificially increase the apparent educational productivity.

There are however three facts that reduce the risk of such a distortion of the measure for educational productivity. Firstly, I can observe a decent set of individual student characteristics which should capture much of the variation across students. Secondly, any unobserved characteristics that can affect test results are likely to have already affected 8th grade ISAT scores and should therefore be at least partially captured in the variable innate. Thirdly, parents in the CPS could not observe educational productivity directly. Accordingly, educational productivity only had a minor effect on school choice decisions (see 8.4. and 8.5.), contrary to the easily observable outcome-based quality measure which had a strong impact on school choice decisions. The fact that educational productivity was not observable and did not drive school choice decisions reduces the risk for the most worrisome potential effect of unobserved student characteristics that could artificially inflate the measure for educational productivity: a systematic accumulation of students with unobservable characteristics that are favourable to high test results.

The introduction of a value-added measure for educational productivity is only likely to have any effects, if this measure provides additional information. If there was a high correlation between observable average outcomes and educational productivity, then there would be no added benefit of introducing the latter.

The correlation between average outcomes and educational productivity of the attended school is rather small at 0.186. The scatter plot below yields additional insights:
Figure 8-2: Scatter-plot of average test results and educational productivity of schools

For most levels of average outcome, schools show a wide spread of educational productivity. Only at the highest levels of average outcome, all schools show a neutral or positive educational productivity measure. But even in the highest range for average outcomes, 4 out of 7 schools show an educational productivity very close to 0.

It is important to note here, that I will standardize the measure for educational productivity to mean 0 and a standard deviation of 1 before I use it in estimations and the simulation. This standardization is necessary to avoid scale effects when comparing the productivity measure to the outcome-based quality measure that was used in the CPS. As the measure for educational productivity is already quite close to mean 0 and a SD of 1, this adjustment is only minor.

8.2. Differentiation of students into types

It is likely, that the basic functions which describe the actions and achievements of students differ systematically across student characteristics. Therefore, I will differentiate the following functions by student type: a) the EPF, b) the choice function for whether a student applies to any school, c) the choice function to determine the attractiveness of each school, d) the prediction of dropout status for those students missing leave-codes of the cohort 2005.
A differentiation of students across more than one dimension strongly increases computational load and programming effort and leads to groups of students that quickly become too small to yield reliable estimates, especially as I also have to differentiate the choice functions by base-years. Therefore, I will only differentiate by one dimension. In order to avoid confusion and likely programming mistakes, I have decided to differentiate all these functions by the same dimension. Potential variables for differentiation are those variables which can be used to clearly differentiate students and which are likely to show differing behaviour across the spectrum of possible observations. Based on these criteria I analyzed the effects of a differentiation by race, lunch status and innate ability.

Lunch status is a proxy for the family income, as the right to receive lunch for free or at reduced prices is pre-conditioned on a low income. However, 87% of the students in the CPS receive reduced-price or free lunch, so that this criterion only singles out the 13% students whose parents are not poor. Moreover, the differences in the various functions across lunch status type are relatively small.

Regarding innate ability, I created three groups of students. High innate students in the quantile covering the 20% of students with the highest 8th grade test scores, low ability students in the quantile covering the 20% of students with the lowest 8th grade test scores and medium ability students. For a differentiation of students by innate ability, there is a problem that arises from systematic attrition from the dataset. For students who dropped out before the PSAE test in 11th grade, I cannot estimate the EPF as I lack the dependent outcome variable. Students who have low test results in the 8th grade ISAT test are more likely to drop out and to miss the PSAE. This is what happened in the CPS, where only 7665 students from the low ability type eventually took the PSAE as compared to 12656 of the high ability type, although both groups originally comprised 20% of those students who started the second year at high-school. This attrition is comparable to a truncation of the lower end of outcome levels (of about 40% of the low-innate group), as students who would have scored very badly on the PSAE drop out before even taking the test. Additionally, students with a low innate ability are unlikely to score very well in the PSAE. The combined effect of a quasi-truncation and limited chances of high scores create an artificially compressed distribution of outcomes for the low innate students. This is reflected in a high intercept and low estimates for the coefficients of the EPF for this group that can be seen in table 8-3. I am therefore hesitant to interpret the low estimated coefficients as an indicator that the corresponding variables have a weaker effect on outcomes for low ability students than for other students.
### Auxiliary EPF by innate ability

<table>
<thead>
<tr>
<th></th>
<th>Low_abil</th>
<th></th>
<th>Medium_abil</th>
<th></th>
<th>High_abil</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>innate</td>
<td>0.583***</td>
<td></td>
<td>0.779***</td>
<td></td>
<td>0.784***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0145)</td>
<td></td>
<td>(0.00715)</td>
<td></td>
<td>(0.0102)</td>
<td></td>
</tr>
<tr>
<td>innate_peer</td>
<td>0.0669***</td>
<td></td>
<td>0.128***</td>
<td></td>
<td>0.226***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0133)</td>
<td></td>
<td>(0.00786)</td>
<td></td>
<td>(0.0136)</td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-0.385***</td>
<td></td>
<td>-0.423***</td>
<td></td>
<td>-1.017***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td></td>
<td>(0.0674)</td>
<td></td>
<td>(0.119)</td>
<td></td>
</tr>
<tr>
<td>lunch_status</td>
<td>-0.272</td>
<td></td>
<td>-0.438***</td>
<td></td>
<td>-3.643***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.224)</td>
<td></td>
<td>(0.112)</td>
<td></td>
<td>(0.154)</td>
<td></td>
</tr>
<tr>
<td>iep</td>
<td>-3.782***</td>
<td></td>
<td>-2.770***</td>
<td></td>
<td>-1.880***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td></td>
<td>(0.135)</td>
<td></td>
<td>(0.559)</td>
<td></td>
</tr>
<tr>
<td>ell</td>
<td>-0.0765</td>
<td>0.836*</td>
<td>-0.398</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.242)</td>
<td>(0.348)</td>
<td>(1.200)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r_African_Am</td>
<td>-0.328</td>
<td>-2.037***</td>
<td>-2.728***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.287)</td>
<td>(0.154)</td>
<td>(0.190)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r_Asian_Pac</td>
<td>-0.0256</td>
<td>1.303***</td>
<td>0.438</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.618)</td>
<td>(0.261)</td>
<td>(0.278)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r_Hispanic</td>
<td>0.635*</td>
<td>-0.996***</td>
<td>-1.394***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.290)</td>
<td>(0.156)</td>
<td>(0.196)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>44.68***</td>
<td>8.591***</td>
<td>-5.105*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.788)</td>
<td>(1.455)</td>
<td>(2.384)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N: 7665  29263  12656
R-sq: 0.384  0.373  0.461
adj. R-sq: 0.383  0.372  0.461
F: 530.4  1930.3  1202.0

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 8-3: Auxiliary educational productivity function by ability level

Overall, the differences were strongest across race, especially for the choice functions. Therefore, I chose race as the dimension by which I differentiate students into types. I will not present the differentiation across all three dimensions here for all estimations. Instead, I will present the differentiation across race in those sections where I present the corresponding estimations (8.3.-8.5 and 8.9.). Differentiations across the other dimensions can be accessed easily by running a do-file for each dimension which is included in the accompanying CD and that I will refer to in each corresponding section.

### 8.3. Estimation of the educational production function

Now that I have computed the educational productivity level for each school and each cohort, I can estimate the real educational production function (EPF), as opposed to the auxiliary EPF before. Table 8-4 holds the estimation results for all students, table 8-5 holds the estimated coefficients for estimations differentiated by race. Note that here only those students can be included for whom I can observe PSAE results, meaning that these students stayed in the area, stayed in public CPS schools, did not drop out and took the test.
### Educational production function

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>innate</td>
<td>0.763***</td>
<td>(0.00345)</td>
<td>female</td>
<td>-0.799***</td>
</tr>
<tr>
<td>innate_peer</td>
<td>0.153***</td>
<td>(0.00683)</td>
<td>lunch_status</td>
<td>-0.622***</td>
</tr>
<tr>
<td>female</td>
<td>-0.799***</td>
<td>(0.0585)</td>
<td>iep</td>
<td>-2.242***</td>
</tr>
<tr>
<td>lunch_status</td>
<td>-0.622***</td>
<td>(0.0923)</td>
<td>ell</td>
<td>0.0243</td>
</tr>
<tr>
<td>elem_repeat</td>
<td>-1.431***</td>
<td>(0.120)</td>
<td>elem_changer</td>
<td>0.264***</td>
</tr>
<tr>
<td>elem_changer</td>
<td>-1.431***</td>
<td>(0.120)</td>
<td>educ_prod</td>
<td>1.223***</td>
</tr>
<tr>
<td>r_African_Am</td>
<td>-3.035***</td>
<td>(0.119)</td>
<td>r_Asian_Pac</td>
<td>0.747***</td>
</tr>
<tr>
<td>r_Hispanic</td>
<td>-1.918***</td>
<td>(0.120)</td>
<td>r_Asian_Pac</td>
<td>0.747***</td>
</tr>
<tr>
<td>r_Asian_Pac</td>
<td>0.747***</td>
<td>(0.120)</td>
<td>r_Hispanic</td>
<td>-1.918***</td>
</tr>
<tr>
<td>elem_repeat</td>
<td>-1.431***</td>
<td>(0.0592)</td>
<td>elem_changer</td>
<td>0.264***</td>
</tr>
<tr>
<td>elem_changer</td>
<td>0.264***</td>
<td>(0.0592)</td>
<td>educ_prod</td>
<td>1.223***</td>
</tr>
</tbody>
</table>

N = 39027
R-sq = 0.753
adj. R-sq = 0.753
F = 9897.9

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

### Table 8-4: Educational production function

Again, innate ability has the strongest impact on outcomes. An increase of 8th grade ISAT test scores by 1 point increases 11th grade PSAE test scores by 0.76 points. An increase of innate ability by 1 SD increases outcomes by 74% of a SD.

An increase of the average 8th grade test scores of the peers (innate_peer) by one point increases outcomes by 0.15 points. This is equivalent to an increase of outcomes by 6% of a SD as a result of an increase of average peer ability by 1 SD.

Boys score better than girls by about 7% of a SD, again probably due to the fact that the composite outcome contains math, science and reading, two subjects where boys usually score better than girls and only one subject where girls usually outperform boys. Poor students (lunch_status=1) score 5% of a SD worse than students who do not receive lunch
subsidies and students in special education score 20% of a SD worse than their peers while the estimate for English Language Learner status is insignificant. African American students score 27% of a SD worse than the reference group, White students. Asian and Pacific Islander students score 7% of a SD better than the reference group and Hispanic students score 17% of a SD worse.

Students who have repeated at least one grade in elementary school (elem_repeat) score 13% of a SD worse than their peers and students who have actively chosen their elementary school (elem_changer) score 2% of a SD better than those who have not.

The educational productivity of the attended school (educ_prod) has a strong impact on outcomes. An increase of 1 point in educational productivity increases outcomes by 1.22 points. This is equivalent to an increase of test scores by 10% of a SD for an increase in the educational productivity by 1 SD. A change from the school and cohort with the lowest educational productivity to the highest productivity in my dataset would result in an increase of scores by 7.2 points or 63% of a SD of outcomes.

The explanatory variables used in this estimation cover 75% of the variation in outcomes. Table 8-5 shows the estimated coefficients for EPF-estimations differentiated by race (The interested reader can find the differentiation by income status or innate ability by running the do-file “C:\Daten\Stata CPS Data\Graphs and Tables\Table for 8.3. EPF by race.do“):

--

43 As mentioned before, my sample only contains those non-native speakers that are judged to speak English well enough to take the ISAT and PSAE test together with the native speakers, instead of taking a special test for non-native speakers. The sample included in my dataset is therefore not representative for non-native speakers.
### Educational production function by race

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th>African_Am</th>
<th>Hispanic</th>
<th>Asian_Pac</th>
</tr>
</thead>
<tbody>
<tr>
<td>innate</td>
<td>0.855***</td>
<td>0.725***</td>
<td>0.772***</td>
<td>0.809***</td>
</tr>
<tr>
<td></td>
<td>(0.0123)</td>
<td>(0.00486)</td>
<td>(0.00560)</td>
<td>(0.0184)</td>
</tr>
<tr>
<td>innate_peer</td>
<td>0.118***</td>
<td>0.163***</td>
<td>0.158***</td>
<td>0.138***</td>
</tr>
<tr>
<td></td>
<td>(0.0305)</td>
<td>(0.00847)</td>
<td>(0.0130)</td>
<td>(0.0389)</td>
</tr>
<tr>
<td>female</td>
<td>-1.707***</td>
<td>-0.296***</td>
<td>-1.168***</td>
<td>-1.760***</td>
</tr>
<tr>
<td></td>
<td>(0.233)</td>
<td>(0.0799)</td>
<td>(0.0946)</td>
<td>(0.333)</td>
</tr>
<tr>
<td>lunch_status</td>
<td>-1.102***</td>
<td>-0.529***</td>
<td>-0.370*</td>
<td>-0.289</td>
</tr>
<tr>
<td></td>
<td>(0.256)</td>
<td>(0.126)</td>
<td>(0.179)</td>
<td>(0.414)</td>
</tr>
<tr>
<td>iep</td>
<td>-2.296***</td>
<td>-2.313***</td>
<td>-2.339***</td>
<td>-3.388***</td>
</tr>
<tr>
<td></td>
<td>(0.388)</td>
<td>(0.132)</td>
<td>(0.172)</td>
<td>(0.739)</td>
</tr>
<tr>
<td>ell</td>
<td>0.192</td>
<td>0.369</td>
<td>0.349</td>
<td>1.401</td>
</tr>
<tr>
<td></td>
<td>(0.815)</td>
<td>(1.748)</td>
<td>(0.274)</td>
<td>(1.078)</td>
</tr>
<tr>
<td>elem_repeat</td>
<td>-1.157***</td>
<td>-1.454***</td>
<td>-1.483***</td>
<td>-1.666***</td>
</tr>
<tr>
<td></td>
<td>(0.237)</td>
<td>(0.0808)</td>
<td>(0.0958)</td>
<td>(0.342)</td>
</tr>
<tr>
<td>elem_changer</td>
<td>0.497+</td>
<td>0.479***</td>
<td>-0.0793</td>
<td>-0.0971</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.0808)</td>
<td>(0.105)</td>
<td>(0.370)</td>
</tr>
<tr>
<td>educ_prod</td>
<td>1.388***</td>
<td>1.167***</td>
<td>1.225***</td>
<td>1.411***</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.0425)</td>
<td>(0.0474)</td>
<td>(0.196)</td>
</tr>
<tr>
<td>_cons</td>
<td>0.303</td>
<td>9.960***</td>
<td>5.127**</td>
<td>5.029</td>
</tr>
<tr>
<td></td>
<td>(4.381)</td>
<td>(1.236)</td>
<td>(1.926)</td>
<td>(5.418)</td>
</tr>
</tbody>
</table>

| N                | 3000        | 20034      | 14704      | 1289       |
| R-sq             | 0.797       | 0.722      | 0.712      | 0.774      |
| adj. R-sq        | 0.796       | 0.722      | 0.712      | 0.772      |
| F                | 1305.1      | 5777.0     | 4032.8     | 485.8      |

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 8-5: Educational production function by race

The most interesting differences in the EPF across race concern the impact of innate ability, average innate peer ability and educational productivity. The groups of students who are usually disadvantaged concerning academic success, African Americans and Hispanics, are most strongly affected by the average innate ability of their peers. Own ability and educational quality of the school still have a strong impact on outcomes for these two groups, but this impact is weaker than for White and Asian/Pacific Islander students. This pattern of impact factors might be due to the fact that most students from the disadvantaged groups attend schools that show very high shares of “minority” students. At a school where most peers show rather low scores, having high scores might not be beneficial for the standing of the student with her peers. Such schools might also lack positive examples of high-achieving students. In such an environment, it is likely that the scores of the peers have a stronger impact on test results relative to the impact of own ability. In contrast to other girls, African
American girls score only slightly worse than boys of the same group. Having chosen the elementary school is only significant for African Americans, but for them, it is significant at the 1% level and associated with a gain in outcomes equivalent to 4.5% of a SD.

8.4. Estimating the parental decision function for: which students choose schools?

In this subchapter I estimate the choice function which will simulate whether a student applies to any school:

\[ P_i^\epsilon (Apply) = X_i \beta^\epsilon + S_{i-1} \gamma^\epsilon + \epsilon_i^\epsilon \]  

\( (1) \)

8.4.1. Explanatory variables for parental decision function 1

As presented in 6.1.1., \( X_i \) is a vector of student characteristics and \( S_{i-1} \) is a vector of characteristics for a previous cohort at the assigned school, that are observable for parents at the time of the school choice decision.

It is important to note here, that I cannot observe intended school choices, but only realized school choice. More precisely, I cannot observe whether a student wanted to attend another school, or whether she did apply to another school. What I can observe is only whether a student attended another school than the assigned home school. The difference between applications for school choice and realized school choice results from the limited number of places. A student might have applied to other schools than the assigned home school but did not win any of the lotteries at the schools she applied to. Therefore, she had to stay at the assigned school. Such a student would show up in my dataset as a non-chooser. However, there is a high density of school in the CPS and most students who applied to any school applied to several schools. Therefore, the probability of being admitted at least at one school of choice should be rather high. And the fact that about half the students in the CPS did not attend the assigned school is an indicator that most of the students who wanted to leave the assigned school managed to do so.

The vector of student characteristics \( X_i \) contains the same variables as for the EPF; innate ability, gender, lunch status, special education status, English Language Learner status, whether the student has chosen elementary school and race dummies.

The vector of school characteristics of the assigned school \( S_{i-1} \) contains the following variables: the share of low-income students, the share of English Language Learner students,
and the share of students of the same race group at the school, the dropout rate and the official quality measure for the school. This quality measure is the share of students who meet or exceed state standards, standardized to mean 0 and a standard deviation of 1. This standardization is necessary, so that the quality measure can be compared to the measure for educational productivity without distortions due to scale- or scope effects. All these school variables contain the observation for the last year that was observable at the time when parents had to choose schools.

As I will simulate school choice decisions based on the estimations of choice behaviour, I can only include those variables in the estimations that I will also be able to generate in the simulation. For some school-level variables, like the truancy rate, I have no student-level observations. The simulation of the school choice process leads to a distribution of students that is different from the one observed in reality. Once students who are truant attend other schools in the simulation than the ones that they attended in reality, the aggregate school-level truancy rates provided by the CPS are no longer valid. As I cannot observe truancy on the student level, I cannot generate truancy levels for schools in the simulation. Hence, I cannot use truancy in the estimations of school choice. Of the school-level variables that were provided by the CPS, I also miss student-level observations for attendance rate and graduation rate. The latter variable would also likely show a high correlation with the dropout rate, which is another reason not to include it in the estimations.

Of the school variables that I could generate from student level characteristics, I did not include the share of special education students. There are three reasons for this exclusion. Firstly, the IEP rate is a rather rough measure. To parents of non-IEP students it is probably more important, whether IEP students are included in the normal classes or taught in special classes and which types of IEP students attend the school. To parents of IEP students, it is probably important, for which types of IEP students the school offers specialized support. This is important, as the label “IEP” is used for students with diverse characteristics, covering for example autism, deaf-blindness, learning disability and emotional disability and orthopaedic impairment (CPS IEP). IEP students in the CPS might therefore be blind or deaf, they might have dyslexia or problems at grasping math, might show behavioural disorders or they might be afflicted by heavy and multiple physical and mental disabilities. A school that has an extensive support program for the blind might not be able to suitably support a dyslexic student. Moreover, in a school that successfully accommodates deaf students in special classrooms, a high IEP rate should not deter parents of non-IEP students. But if the same high IEP rate is explained by students with behavioural disorders who are integrated in
normal classrooms, many parents will not want their child to apply to this school. The second
reason for not including the IEP share is, that this characteristic was not made easily
accessible to parents on the school report cards or on the internet. I had to compute it myself
based on student-level data. The third reason for not including the IEP share is, that IEP
students scored considerably worse in the 11th grade PSAE test (by 20% of a SD). Therefore,
an increase in the share of PSAE students decreases the quality measure that is also included
in the estimations of choice behaviour. Auto-correlation effects would therefore likely distort
the impact of the quality measure that was observable to parents if I also included IEP shares,
that were not observable to parents.

8.4.2. Estimation of parental decision function 1

I estimate the decision function for whether a student applies to any school with a logistic
binary estimation with a variable on whether the student attended another than the assigned
school as the dependent variable. This variable is called “hs_changer” and is equal to 1 if the
student did not attend the assigned school. As the output of binary estimations is hard to
interpret, I use the command “listcoef, std” which gives the impact of each variable measured
in standard deviations of the dependent variables. For binary explanatory variables, this
command returns the change in probability, as measured in SD of hs_changer, for the
difference between 0 and 1, in the column named bStdY. For continuous variables, this
command returns the impact of a 1 SD increase in the explanatory variable as measured in SD
of hs_changer, in the column named bStdXY.
Note, that the SD of the explanatory variable is 0.499. Therefore, if the estimates are divided by two, this results in the change of the probability to actively choose a high-school that is associated with the explanatory variable. A positive value means, that the probability that a student tries to leave the school increases, if the corresponding explanatory variable increases.

I observe the following findings concerning student characteristics: Students who had actively chosen elementary school (elem_changer) show a probability of not attending their assigned high-school that is 0.13 higher than for their peers. Meaning that if a student with otherwise identical characteristics had a probability of 0.5 not to attend the assigned school, a student who in addition had already chosen her elementary school would have a probability of 0.63 not to attend the assigned school.

For girls (female), the probability to choose a school is 0.05 higher, for poor students (lunch_status=1) 0.04 lower, for IEP students 0.04 higher and for English Language Learner (ell) students 0.06 higher than for their peers. The probability for African American students not to attend the assigned school is 0.18 higher than for the reference group, White students. For Asian and Pacific Islander students, the probability is 0.09 higher and the estimate for Hispanics is not significant.

The variables on innate ability and observed school quality have the strongest impact on the probability of choosing a school. An increase in innate ability by 1 SD increases the probability not to attend the assigned school by 0.15, meaning that high ability students are more likely to apply to other schools. An increase in the observed quality of the assigned
school (quality_ass) by 1 SD decreases the probability of attending another school by 0.17, meaning that the better the school scored on the official quality measure, the less likely its students are to apply to other schools.

Concerning other aggregate school characteristics, the probability to leave the assigned school increases with the dropout-rate and decreases with the share of students of the own race and the share of ell students at the assigned school. The estimate for the share of students from low-income households is not significant.

The following table holds the “listcoef, std” results differentiated by race (The interested reader can find the differentiation by income status or innate ability by running the do-file “C:\Daten\Stata CPS Data\Graphs and Tables\Table for 8.4. differentiate estimation for school choice or not.do”. This file also includes the outcomes of the logit-estimations for the estimation for all students and differentiated by race):

<table>
<thead>
<tr>
<th>Race group</th>
<th>White</th>
<th>African_Am</th>
<th>Asian_Pac</th>
<th>Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bStdY</td>
<td>bStdXY</td>
<td>bStdY</td>
<td>bStdXY</td>
</tr>
<tr>
<td>elem_changer</td>
<td>0.2777  ***</td>
<td>0.2388 ***</td>
<td>0.2299 ***</td>
<td>0.3520 ***</td>
</tr>
<tr>
<td>female</td>
<td>0.1509  ***</td>
<td>0.1049 ***</td>
<td>0.1121 **</td>
<td>0.0792 ***</td>
</tr>
<tr>
<td>lunch_status</td>
<td>-0.2419 **</td>
<td>-0.0077</td>
<td>-0.2238 ***</td>
<td>-0.1750 ***</td>
</tr>
<tr>
<td>lerp</td>
<td>0.0536</td>
<td>0.0656 ***</td>
<td>0.0302</td>
<td>0.1719 ***</td>
</tr>
<tr>
<td>ell</td>
<td>0.2353  **</td>
<td>0.2915</td>
<td>0.1945</td>
<td>0.0287</td>
</tr>
<tr>
<td>innate</td>
<td>-0.3215 ***</td>
<td>0.2994</td>
<td>0.3770 ***</td>
<td>0.2823 ***</td>
</tr>
<tr>
<td>quality_ass</td>
<td>-0.2721 ***</td>
<td>-0.3905 ***</td>
<td>-0.2204 ***</td>
<td>-0.2949 ***</td>
</tr>
<tr>
<td>lowin_ass</td>
<td>-0.0509 ***</td>
<td>0.0495 ***</td>
<td>0.2227 ***</td>
<td>0.0532 ***</td>
</tr>
<tr>
<td>dropout_ass</td>
<td>-0.3205 ***</td>
<td>-0.0375 ***</td>
<td>0.1119 ***</td>
<td>-0.2718 ***</td>
</tr>
<tr>
<td>own_race_ass</td>
<td>0.0598</td>
<td>-0.0113 ***</td>
<td>0.2556 ***</td>
<td>-0.0014 ***</td>
</tr>
</tbody>
</table>

-Not significant *significant at 10% level **5% level ***1% level

Table 8-7: Parental decision function 1 by race

The following differences are interesting. For White and Hispanic students the probability to leave the assigned school decreases strongly with the share of students of the own race, while for African Americans this share has only a minor impact and for Asian and Pacific Islander students the probability to leave the assigned school increases when the share of their own group increases. African American students react most strongly to the observed quality of their assigned school, the change in probability is almost twice as strong as for Asian students and much stronger than for White and Hispanic students. The impact of innate ability on the probability to actively choose a school is strongest for Asian students, weaker for White students and weakest for African American and Hispanic students.

The simulated choice behaviour will be based on estimations that are differentiated by race and the base-year which was used for the estimation.
8.4.3. Cut-off values for the predictions based on the parental decision function 1

For the simulation, I will predict the probability to apply to any school based on the estimations described above. In order to make predictions of binary variables such as whether a student applies to any school or not, I cannot simply use predictions of these estimations. There is an additional step that needs to be taken. The prediction results in the probability with which the student applies to any school. This probability is a value between 0 and 1. In order to generate a binary variable that determines, which students will apply to schools, I need to define a cut-off value. All students who show a predicted probability for choosing a school that is above the defined threshold will be declared choosers and will apply to other schools.

These cut-off values have to be chosen carefully as they affect the characteristics of the average school-changer and therefore most outcomes of the simulation. The effects of different levels of the cut-off level can best be illustrated by using the example of the variable with the strongest impact on choice behaviour: innate ability. Students with a high innate ability are more likely to choose schools. Therefore, they get assigned a higher predicted probability to apply to any school. Consequently, the average school-chooser has an innate ability that is above the average of the entire population. At the same time, the high innate ability results in high test scores. Thus, a school that attracts many school changers attracts many students with an innate ability above the average and will therefore have high average test results when these students take the PSAE. If the cut-off value increases, the share of choosers is reduced, but their average innate ability increases and schools which attract many choosers achieve even better average test results.

Thus, a change in the cut-off value affects the characteristics of choosers, the share of choosers, school outcomes and sorting by ability. By affecting the composition of the group of students who choose schools, the cut-off value also affects the effects of actively choosing a school.

I calibrated the cut-off values by incremental adjustments until a simulation that was designed to be as close to the reality as possible yielded an average share of choosers that was as close as possible to the share that was observed in reality in the CPS. As choice behaviour differs across race, I had to calibrate that cut-off value for each student group separately. And as I will randomly use estimations of different base years for robustness reasons, I had to repeat the entire process for each base year.
8.5. Estimating the parental decision function for: which schools are chosen?

Based on the estimation results described in 8.4. I can predict which students will apply to any school rather than just stay at the assigned one. In this subchapter I present the estimations that I will use to predict preferences for other schools. The parental utility function for school \( J \) for student \( I \) of type \( X \) is, as presented in 6.1.1.:

\[
U_{ij}^x = f(D_{ij})^* (X_i \beta_x^U + S_{ij}^{-1} \gamma_x^U + S_{ij}^{-1} \delta_x^U + \varepsilon_{ij}^U) \quad (2),
\]

The first part that covers the impact of the distance between the home of the student and the school will be covered in the next subchapter, 8.6. In this subchapter I only estimate the second part of the utility function. Note that this estimation does not result in a utility value in the classic sense, but in the probability of student \( I \) of type \( X \) to attend school \( J \) if given a choice between this school and the assigned school \( I \).

8.5.1. Explanatory variables for parental decision function 2

Again, I do not observe applications to schools but only realized changes. When estimating the preferences for school characteristics, this might be a more serious problem than in the subchapter before. Concerning the estimation about whether students apply to any school, I could safely assume that most students who do not want to stay at their assigned school get access to any school of their choice and that I can therefore identify them as choosers. Here, I cannot assume that each chooser got access to the school of her highest preference. If a student only got access to her third or fourth choice, those characteristics of the attended school that made it attractive to the student will show lower values than the unobservable first choice. These pull-characteristics will also likely show different weights. The differences between characteristics of the unobserved first choice and characteristics of the attended school distort the preferences of students. An illustrative example might clarify the nature of this distortion: a student cares very much for high test scores and a little less for dropout-rates of schools. She does not get access to her first two choices with very high test scores, but only to her third choice, which has a very low dropout-rate and rather high test scores. As I can only observe the realized third choice, this student will appear to care relatively more about dropout-rates and less about test scores than she does in reality.

However, students will only change to schools that match their preferences better than the assigned school. And the mechanism described above might also work in the other direction. If the third-choice school has rather high test scores but only a mediocre dropout-rate, the
student would appear to care more about test scores relative to the dropout-rate than she really does. Therefore, averaged over a bigger group of students, distortions of relative preferences that arise from not getting access to the first choice school are likely to cancel each other out. At least, the average chosen school will show elevated variable values for those characteristics that students like and lower values for characteristics that students dislike. I include in this estimation the same variables as for the estimation on whether a student applies to any school and add the school characteristics for the attended school. Thus, the variables of the estimation comprise individual student characteristics and school characteristics for the assigned and for the attended school. The inclusion of characteristics of the assigned school allows me to observe and analyze push-factors, which make a student want to leave an assigned school, together with the pull-factors that make a student want to apply to another school.

8.5.2. Estimation of parental decision function 2

I estimate the decision function for attractiveness of schools using a logistic binary estimation with a variable on whether the student attended another than the assigned school as the dependent variable. This approach is not ideal in order to identify school preferences. In the best conceivable setting, I could observe all schools to which a student applied in the order of her preference. By comparing the characteristics of chosen schools, the relative impact of school characteristics could be estimated precisely. Unfortunately, the CPS could not provide this information, mostly due to the fact that most schools handle their own application procedure. Thus, only the individual schools have information on which students applied there. Moreover, application information was not stored in a standardized way, so that an aggregation by CPS staff would have been prohibitively time-consuming and hence prohibitively expensive. And privacy protection rules made it impossible for anyone but CPS staff to get access to application records from individual schools. Therefore, I have to use characteristics of the attended schools in order to estimate school preferences.

As my dependent variable is a binary variable on whether the student attended another school than the assigned home school, I have to include all students in the estimation instead of focussing on the school-choosers only. Otherwise, there would be no variation in the dependent variable. Before I start to present the estimated values, note that, given the available data, these estimates should not be interpreted as real changes in the probability to apply to the
alternative school. Based on the estimation approach, these estimates show the impact of the variables for attending the school, if given the choice between this school and the assigned school. Accordingly, I will not use these estimates to directly predict the probability of applying to individual schools, but to approximate the preferences of parents. Then I will only use predictions based on the above estimation for the pair wise comparison of predicted probabilities for schools. If, for example, the predicted probability of choosing school A instead of the assigned school is higher than the predicted probability of choosing school B instead of the assigned school, I will deduce that school A is more desirable to the student than school B.

As the output of binary estimations is hard to interpret, I use again the command “listcoef, std” which gives the impact of each variable as measured in standard deviations of the dependent variables.44

logit (N=61913): Unstandardized and Standardized Estimates

|            | b     | z     | P>|z| | bStdX | bStdY | bStdXY | SDofX |
|------------|-------|-------|------|-------|-------|--------|-------|
| elem_changer | 0.46829 | 20.732 | 0.000 | 0.2185 | 0.1640 | 0.0765 | 0.4666 |
| female      | 0.14986 | 7.118  | 0.000 | 0.0749 | 0.0525 | 0.0262 | 0.5000 |
| lunch_status| 0.07440 | 2.078  | 0.038 | 0.0245 | 0.0261 | 0.0086 | 0.3287 |
| iep         | -0.06482 | -1.974 | 0.048 | -0.0247 | -0.0227 | -0.0087 | 0.3816 |
| ell         | 0.20215 | 2.381  | 0.017 | 0.0278 | 0.0708 | 0.0097 | 0.1374 |
| innate      | 0.02515 | 19.837 | 0.000 | 0.2835 | 0.0088 | 0.0993 | 11.2748 |
| r_Asian_Pac | 0.76736 | 15.057 | 0.000 | 0.3832 | 0.2687 | 0.1342 | 0.4994 |
| r_Hispanic  | -0.05072 | -1.039 | 0.299 | -0.0245 | -0.0178 | -0.0086 | 0.4831 |
| quality_ass | -1.94073 | -56.909 | 0.000 | -1.4833 | -0.6818 | -0.5194 | 0.7618 |
| quality_att | 1.37838 | 42.464 | 0.000 | 1.3714 | 0.4827 | 0.4802 | 0.9949 |
| lowin_ass   | -0.01168 | -6.342 | 0.000 | -0.1295 | -0.0041 | -0.0453 | 11.0855 |
| lowin_att   | 0.02064 | 12.144 | 0.000 | 0.2851 | 0.0072 | 0.0998 | 13.8180 |
| dropout_ass | 0.11156 | 44.288 | 0.000 | 0.7790 | 0.0391 | 0.2728 | 6.9823 |
| dropout_att | -0.14902 | -60.940 | 0.000 | -1.0601 | -0.0522 | -0.3712 | 7.1139 |
| own_race_ass| -0.01335 | -21.834 | 0.000 | -0.4287 | -0.0047 | -0.1501 | 32.1049 |
| own_race_att| 0.00676 | 10.539 | 0.000 | 0.2159 | 0.0024 | 0.0756 | 31.9210 |
| ell_ass     | 0.12971 | 6.136  | 0.000 | 0.1393 | 0.0454 | 0.0488 | 1.0739 |
| ell_att     | -0.42208 | -18.864 | 0.000 | -0.4361 | -0.1478 | -0.1527 | 1.0332 |

Table 8-8: Parental decision function 2: Which school?

Again, the SD of the dependent variable is close to 0.5, so that dividing the values in bStdY or bStdXY by two provides the change in the probability of being a high-school-chooser that results from changes in the explanatory variable.

44 The “listcoef, std” command and how the results are to be interpreted are explained in 8.4.2.
A student who had actively chosen her elementary school has a probability of choosing high-
school that is 0.08 higher than that of her peers. The probability to choose is 0.026 higher for
girls, 0.013 higher for poor students, 0.011 lower for IEP students and 0.035 higher for ELL
students than for their peers. The probability not to attend the assigned school is for African
American students 0.134 higher than for the reference group, White students. The estimate for
Hispanic students and Asian/Pacific Islander students are not significant. An increase in
innate ability of 1 SD increases the probability of choosing by 0.05.
The strongest impact on school choice decisions have the variables on observed school
quality and the dropout-ratio. An increase in the latest observable measure for academic
quality by 1 SD in the assigned school decreases the probability to choose another school by
0.26, while an increase of 1 SD in the quality measure of the alternative school increases the
probability to choose by 0.24. A 1 SD increase in the dropout rate of the assigned school
increases the probability of leaving this school by 0.136, a 1 SD increase of the dropout rate
for the alternative school decreases the probability of choosing this school by 0.175.
Students prefer schools where a high share of the student population shares their race. The
impact of the share of the own race is stronger in keeping students at the assigned school than
as a sought-after characteristic for a chosen school. Students dislike a high share of English
Language Learners. Here, the pull-factor is stronger than the push-factor. A high share of ell
students only slightly pushes students to leave the assigned school, but a 1 SD increase in the
ell-share of the alternative school decreases the probability of choosing this school by 0.076.
Concerning the low-income rate, students seem to prefer schools with a high share of low-
income students. A 1 SD increase in the share of low-income students at the alternative school
increases the probability of choosing this school by 0.05. This might be a hidden impact of
distance. The share of low-income students in the CPS is above 85%. Most schools in the
CPS therefore have a high low-income share.
Consequently, most students live in areas, where all schools that are not far from the home of the student have a high low-income share. The few schools with a low low-income share tend to be in wealthier areas and their students are less likely to change schools. Thus, for a large share of potential choosers, the low-income share might be a proxy for distance and a high share of low-income students might therefore appear to be attractive to students.

As for the other estimations described above, I differentiate by race (the interested reader can find the differentiation by income status or innate ability by running the do-file “C:\Daten\Stata CPS Data\Graphs and Tables\Table for 8.5. differentiate estimation for which school.do“):
Table 8-9: Parental decision function 2 by race

The following differences are interesting: Whether the student had actively chosen elementary school has the strongest impact for students of those groups that are usually academically disadvantaged, African American and Hispanic students. For African American students, innate ability has a much stronger impact on the probability to choose the alternative school than for the other groups. The impact of the observed academic quality of the alternative school has a much lower impact on the probability to choose the alternative school for African American students, while the impact of the assigned school does not differ that strongly to the impact for other groups. Finally, the dropout rate has by far the highest impact of all groups. According to these observations and the ones on whether to choose a school at all, African American students are more likely to leave the assigned school if their innate ability is higher. These students leave schools with a weak quality measure and a high dropout ratio and when choosing another school, they care very strongly about the dropout rate, but less strongly about the quality measure of these schools. Students of the other groups also leave low quality and high drop-out schools but when they choose, they mainly pursue high quality measures.

Another observation is, that White students are hesitant to leave schools where a high share of the students belongs to the same group while Asian/Pacific Islander students seem to dislike a high share of students of their own group.

The strong differences by race in the observed school choice behaviour, both in chapter 8.4. and here, strongly drove my decision to use race as the variable by which I differentiate students into groups.
8.5.3. Predictions for use in the simulation based on parental decision function 2

Based on the estimation results above, I can predict the probability of each student to attend each alternative school, if given the choice between this school and the assigned school. Contrary to the predictions for whether a student applies to any school, I do not need cut-off values, as I can compare the probabilities directly and treat them as utility levels. If the predicted and distance-adjusted probability to apply to school $X$ is higher than the probability to apply to school $Y$, then the student prefers school $X$ over school $Y$.

8.6. Modelling the impact of distance

As described above in 6.2.1.(4), I do not have information on the home location for individual students. Instead, I have to approximate the home location of each student by the location of her assigned home school. This proxy is imprecise, especially if students only travel short distances to their school of choice. And most students only travel distances where this imprecision of the proxy is a major issue. Moreover, the distance-proxy is necessarily equal to zero for those students who attend their assigned school. Therefore, I do not estimate the impact of distance on school choice directly.

But a higher distance to school is associated with unfavourable effects. Consequently, the distance to school has been shown to be an important factor in the school choice decision (see 3.4.2.(3) and 6.2.1.). Therefore, I have decided to include the impact of distance by discounting the attractiveness of potential schools that is based on other characteristics by a distance-factor. The reasoning behind this approach is as follows:

Distance lowers the attractiveness of a school. It does not affect other favourable characteristics directly, like good test scores in previous years. But distance is associated with costs such as the time and money spent on the daily commute and a higher probability of being separated from friends and classmates. I use a discounting approach to represent how the costs of distance lower the attractiveness of a school in the same way that future income is assumed to be considered less valuable than income today. The multiplicative nature of a discounting factor has, as opposed to a summary impact of distance on the overall attractiveness of a school, the important advantage of being immune to two types of scale effects.

A summary factor to represent the impact of distance would have a fixed total size that depends only on the distances. The total size of the number that represents the attractiveness of other schools depends on their quality as perceived by parents. In an area with very
attractive schools, a summary distance factor would have a low impact relative to other characteristics. In an area with rather unattractive schools, a summary distance factor of the same size would have a much stronger impact relative to other characteristics. This is the first type of scale effect that can be avoided by using a multiplicative approach to the impact of distance.

The second type of scale effect occurs for example, when competition has an impact on overall school productivity. If the productivity of all schools rises, then the total size of the number that represents the attractiveness of a school based on all characteristics except distance increases and differences in distance have a lower impact on school choice decisions. If the impact of distance is modelled in a multiplicative way, like a discounting factor, its effect is independent of the total size of the number that represents the attractiveness of a school based on other characteristics.

However, a multiplicative factor is associated with another potential problem. If this factor is allowed to vary too strongly, it can overwhelm the impact of all other variables. If, as for example in an earlier version of the simulation, the discount factor varies between 0 and 1, school choice is almost exclusively driven by distance to potential schools.

Students are not willing to travel indefinitely far each day in order to attend an attractive school. I therefore added a maximum distance to schools at a distance of 24 kilometres. Schools that are farther away are not considered by students in the school choice process of my simulation. This cut-off value is based on the distribution of distance to the attended school that can be seen below in figure 8-4. Only 1% of those students in the CPS who do not attend their home school attend a school that is more than 24 kilometres distant from their assigned school.

Based on the above considerations, the shape of the function for the impact of distance is:

\[ f(D_j) = B + \frac{a^\max - D_j}{a^\max} \]  

As I cannot estimate the impact of distance due to a lack of precise data, I have used a trial and error approach, first varying the shape of the function for the impact of distance and then slowly varying the value of B until the resulting distribution of distances to the attended school in my simulation closely resembles the distribution observed in reality. This calibration was complicated by the fact that I also had to calibrate the cut-off values above which students would apply to any school at all (as described in 8.4.). The variables concerning the impact of distance and the propensity to apply to schools affect each other. If I increase the impact of distance for example, students are less likely to prefer a distant school
to their assigned school. Therefore, in order to retain the rate of choosers as observed in the real world data, I would have to lower the cut-off values for choice.

As I calibrated cut-off values for each base-year and race separately, these calibrations were rather time-consuming. The cut-off values need to be calibrated carefully, as explained in 8.4, as they affect the average characteristics of school-changers and thus all relevant outcomes of the simulation. Consequently, I focused in the calibration more on getting the real world school-chooser shares close to the real world values than on getting the distance-distribution close to the observed distribution. The value for $B$ that resulted from the described calibration method is 2.5.

The figures below illustrate how close the distribution of distances that results from the simulation is to the distribution of distances in reality. It is important to note here, that this is a comparison to a simulation outcome that was generated with the use of randomizations. It can therefore not be reproduced exactly. However, the pattern shown in this example is persistent.

![Figure 8-4: Distance to the attended school, dataset](image.png)

45 For the simulation that I describe and use later on, I use randomizations of several variables as a robustness measure (see 8.10). Any randomized simulation run will have differing variables and result in a slightly different distribution of distances. In the accompanying CD, in the file “Daten\Stata CPS Data\Graphs and Tables\Graph distribution after 4 turns X.do” can be found two examples for datasets that result from a simulation. The table can be generated with the command “hist distto_attended if distto_attended!=0” when these datasets are loaded. More datasets can be generated by running: "Daten\Stata CPS Data\A Master for distance distribution.do" and setting the number of turns in line 111 of "C:\Daten\Stata CPS Data\Simulation including Schoolmax\Choice and Outcome Generation Schoolmax.do" to the value 4. The resulting dataset can be found under "Daten\Daten CPS\Dataset Simulation Outcomes Out Schoolmax.dta"
8.7. Competitive pressure

In this subchapter, I describe the computation of the measure for competitive pressure by other schools. As presented in chapter 6.1.1., the function for competitive pressure is:

\[ C_{j} = \sum_{n \neq j} f(D_{nj}Q_{n}^{-1}) \quad (6), \]

Competitive pressure increases with the number of nearby schools, it increases with the academic quality of individual nearby schools and it decreases with the distance to individual competing schools.

8.7.1. The impact of relative academic quality on competitive pressure

The higher the academic quality of a competing school, the higher is the competitive threat that this school poses. In the CPS, the official measure for academic quality is the share of students who meet or exceed state standards. To measure the impact of academic quality, I use quality relative to the average quality in the school choice system.

To an individual school, it would be more important, whether a competing school shows a higher academic quality than the own school, not whether the competitor has a higher quality than the average quality. However, if I used the quality of the own school in estimating educational productivity for the same school, I would face reverse causality problems. Educational productivity is highly correlated across time. Additionally, a higher educational productivity results in higher scores, and thus in a higher academic quality measure. Thus, if I included the own academic quality of a school in a function explaining educational
productivity, there would be a channel for a strong effect of the dependent variable on an explanatory variable, resulting in reverse causality. Therefore, I measure the academic quality of competing schools relative to the overall mean, not relative to the quality of the individual school for which I am computing the measure for competitive pressure. As I have previously standardized the quality measure to mean zero, the academic quality relative to the mean is for each school identical to the current measure for academic quality of this school.

During the simulation, I will always use the starting mean of zero, instead of taking a new mean each turn based on the values present in the current turn. The reason is, that I do not want to negate effects of overall increases in educational productivity. If, as a reaction to an overall increase in competitive pressure, all schools increase their educational productivity, the academic quality measure of all schools increases as a result. Then schools would have to maintain a higher level of productivity in order not to lose competitiveness. If I would use the academic quality of other schools relative to the new mean in computing the competitive threat to schools, the overall increase in academic quality of the last turn would be negated. Consequently, the fact that all competing schools had increased their educational productivity in the last turn would have no effect on the effort, and thus on the level of productivity of an individual school in the following turns.

In order to determine how strong the effect of the relative academic quality of competing schools is on competitive pressure, I have analyzed change-patterns concerning quality. That is, I have analyzed the distribution of changes in academic quality that resulted from observed school choices in the CPS. These changes reflect the preferences of the student body for school quality and can thus be used to approximate the threat that a competing school poses based on its quality level. The graph on the next page shows the distribution of the apparent gain in academic quality from not attending the assigned school. This apparent gain is based on the quality measure that was used in the CPS, the share of students at the school who met or exceeded state standards, with the values that were observable when the school choice decision was made.
Most students who actively choose schools, attend those schools which show a higher academic quality than the assigned school. Only 2.8% of the choosers attend a school with a quality more than 1 SD below that of the assigned school, 4.7% attend a school with a quality between 0.5 and 1 SD below that of the assigned school and 13.4% of changers attend a school with a quality between 0.5 and 0 SD below that of their assigned school. Most choosers attend a school with a higher academic quality than their assigned school. 22% of changers gain 0 to 0.5 SD, another 18.8% gain 0.5 to 1 SD and 38.3% gain more than 1 SD of academic quality as a result of leaving their assigned school for another one.

I use this distribution of changes in quality as a starting point to model weights for the competitive pressure that is posed by individual schools, based on the relative academic quality of these schools. As presented above I will have to use academic quality relative to average quality in the simulation, as opposed to academic quality relative to the quality of the individual school. I will therefore use academic quality relative to average quality from now on.

20% of school changers attend a school with an academic quality below average. Thus, I assign 0.2 as the average quality-weight for schools with a quality below average quality. Hardly any student attends a school with a quality more than 1 SD below that of the assigned school. I therefore choose to consider schools with an academic quality measure more than 1
SD below average as posing no threat, resulting in a lower bound of -1. The function for the quality-weight for school $n$ with a quality below average is therefore 

$$W_n^Q = 0.4 \times [1 + (Q_n^{i-1} - 0)]^{46}$$

if the academic quality of school $(Q_n^{i-1})$ is below average. Note that the relative academic quality is based on the newest observable value at the time when parents have to make school choice decisions. 80% of changers choose a school with an academic quality above average. Thus, I assign the average quality-weight of 0.8 to schools with a quality above average. The weight should increase with the quality of the competing school, but it is not likely that the potential threat of a school increases indefinitely with relative quality. A competing school that has an academic quality measure 2 SD above that of the own school is clearly much better and will attract applications from most students, whose parents care about quality and who are actively choosing schools. Moreover, a school with a quality measure 2 SD above the own school will likely receive more applications than it has places for students. A further increase in quality might lead to more applications, but the school would not be able to take in more students. Therefore, whether this school even has an academic quality 3, instead of 2, SD above that of the own school is not likely to increase the competitive pressure much further, and definitely not by the same amount than the difference between 0 and 1 SD of relative quality advantage. Also, even for a rather low-quality school with an academic quality 1 SD below average, an increase by 2 SD covers 83% of those schools that have a higher academic quality. Moreover, about 80% of choosers show a quality-gain of less than 2 SD. Finally, although it is realistic that schools with a higher quality pose a higher threat, I do not want the difference in the threat level of a very good school relative to an average school to become so big that it overwhelms any impact of other characteristics of schools. The reason is, that a close-by school could attract students with other characteristics than a higher academic quality measure. Examples would be a special focus on certain subjects, sports, the arts, communication skills or the simple fact that siblings already attend the school.

Based on the above considerations, I set an upper bound on the quality-weight at the value for an academic quality 2 SD above average. In order to achieve an average weight of 0.8 for schools with an academic quality above the average, I can again use the weight-function

---

46 $(Q_n^{i-1})$ is in the range between -1 and 0 for schools with a quality below average. The value in the brackets is then between 0 and 1. The middle point is 0.5 for a value of $(Q_n^{i-1}) = -0.5$. $(0.4 \times 0.5) = 0.2$, hence $0.4 \times [1+...]$
defined for schools with an academic quality below average, $W_{n}^{Q} = 0.4 * [1 + (Q_n^{r-1} - 0)]$ \(^{47}\). The resulting complete function for the quality-weight is:

$$W_{n}^{Q} = \begin{cases} 
0 & Q_n^{r-1} < -1 \\
0.4 * [1 + (Q_n^{r-1} - 0)] & Q_n^{r-1} \in [-1, 2] \\
1.2 & Q_n^{r-1} > 2 
\end{cases} \quad (7)$$

Based on this function, the best schools show a quality-weight for the computation of competitive threat of 1.2 that is 3 times as high as that of an average quality school, at 0.4.

### 8.7.2. The impact of distance on competitive pressure

Figure 8-4 already showed the distribution of school-choosers in the CPS by the distance the students travelled to the attended school. In the CPS, the average distance for a school to the next school is 2.3 km. For 76.3% of the students, the school closest to their assigned school is at a distance of 3 km or less. For the schools of 94.5% of the students, there is one competitor within a distance of up to 4 km.

The graph below shows (on the line marked with diamonds) which percentage of students (x-axis) travelled as far or farther than the corresponding distance (y-axis). I use this percentage as a weight that I assign to the competitive pressure of a competing school based on the distance. The line marked with squares shows a function that closely approximates the distribution that was observed in reality. In the simulation, I will use the function that underlies this line to weigh the competitive threat posed by other schools based on distance.

\(^{47}\) ($Q_n^{r-1}$) is in the range between 0 and 2 for schools with a quality below average. The value in the brackets is then between 0 and 2. The middle point is 2 for a value of ($Q_n^{r-1}$) = 1. (0.4*2)=0.8, hence 0.4*[1+…]
The underlying function for the distance-weight of school $n$ when competing with school $j$ and based on the distance between these two schools ($D_{nj}$) is:

$$W_{nj}^D = \begin{cases} 
(1 - 0.082 * D_{nj}) & D_{nj} \leq 10 \\
0.18 - 0.002667 * D_{nj} & D_{nj} \in [10, 16] \\
0 & D_{nj} > 16 
\end{cases}$$

(8)

8.7.3. The measure for competitive pressure

Using the weight-functions for distance and relative academic quality, I can now construct the function for competitive pressure faced by school $J$ and based on the distance to and relative quality of competing schools:

$$C_{j}^{r-1} = \sum_{n \neq j} f(D_{nj}, Q_{n}^{r-1}) = \sum_{n \neq j} W_{nj}^{D} * W_{n}^{Q}$$

(9)

The competitive pressure based on this function increases with the number, closeness and relative academic quality of competing schools.
8.8. The productivity determination function

In this subchapter, I estimate the function that determines the educational productivity that is provided by schools.

8.8.1. Estimation of the productivity determination function

As presented in 6.1.1., educational productivity is affected by the capacity of the school that remained free in the last turn, by competitive pressure and by probation status. The corresponding function is:

\[ P_j = F_j^{r-1} \theta^r + (F_j^{r-1})^2 \theta^w + C_j^{r-1} \eta^r + (C_j^{r-1})^2 \eta^w + p_j' + \epsilon_j^P \] (5)

The OLS-estimation of the above function results in:

<table>
<thead>
<tr>
<th>Productivity determination function</th>
<th>educ_prod</th>
</tr>
</thead>
<tbody>
<tr>
<td>competition</td>
<td>1.427**</td>
</tr>
<tr>
<td></td>
<td>(0.537)</td>
</tr>
<tr>
<td>compet_sq</td>
<td>-0.188*</td>
</tr>
<tr>
<td></td>
<td>(0.0864)</td>
</tr>
<tr>
<td>freecapacity</td>
<td>2.236+</td>
</tr>
<tr>
<td></td>
<td>(1.285)</td>
</tr>
<tr>
<td>freecap_sq</td>
<td>-2.186</td>
</tr>
<tr>
<td></td>
<td>(1.871)</td>
</tr>
<tr>
<td>probation</td>
<td>-0.641***</td>
</tr>
<tr>
<td></td>
<td>(0.189)</td>
</tr>
<tr>
<td>_cons</td>
<td>-2.397**</td>
</tr>
<tr>
<td></td>
<td>(0.797)</td>
</tr>
</tbody>
</table>

| N                                  | 216       |
| R-sq                               | 0.120     |
| adj. R-sq                          | 0.099     |
| F                                  | 5.718     |

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 8-10: Estimation of the productivity determination function

This estimation yields the following insights: an increase in the competitive pressure faced by other schools increases the educational productivity, but with decreasing returns. Due to the decreasing returns, the impact of a change in competitive pressure changes with the absolute value of competitive pressure. An increase of competitive pressure by 1 SD, from 0.5 SD below the average to 0.5 SD above the average, results in an increase of educational productivity by 14% of a SD. An increase in last year’s free capacity also increases the educational productivity of a school, also with decreasing returns. An increase of free capacity by 1 SD, from 0.5 SD below the average to 0.5 SD above the average, results in an increase of educational productivity by 17% of a SD. The effect of being on probation does not show the expected sign: being on probation in the last turn results in a lower level of educational
productivity. There might be two effects at work here. The first possible effect is that those teachers who can do so leave a school that is on probation. As better teachers are more likely to find another employment, this might drain the educational productivity of a school. The second possible effect is, that schools under probation might resort to cheating, as this might be an easier way to get out of probation. The performance policy measure, which is used to judge whether a school is set on probation, relies heavily on test outcomes of students and does not consider the educational productivity. As an increase in productivity only indirectly affects outcomes and requires much effort, schools under probation that need to make fast progress as measured in test outcomes, are likely to rely more on cheating measures. Examples for such cheating measures are teaching to the test, teaching “the” test, focussing on the students that are close to meeting state standards or going to such extremes as “correcting” student responses in the tests or getting rid of the weakest students by increasing the drop-out rate.

The decreasing returns of the effect of free capacity and competitive pressure result in a peak within the range of observed values, as can be seen from the following graphs:

![Graph of the effect of competitive pressure on educational productivity](image)

**Figure 8-8: Effect of competitive pressure on educational productivity**

The impact of competitive pressure peaks at a value of 3.8, which is above the average value for competitive pressure that was observed in the CPS at 3.3 and well within the observed range of [0.73, 5.00]. As mentioned in 6.2.2.(3) the decreasing returns might be due to three effects; a) the increased costs of further increasing productivity from a high level, b) the possibility of using easier ways to increase productivity (cheating) that becomes more attractive relative to productivity the higher productivity is and the more effort therefore is
needed for a further increase and c) the possibility that those teacher who can do so flee a failing school and the remainder might cease effort in the face of overpowering competition.

The impact of free capacity peaks at 51%, which is well above the observed average of 33%, and starts to fall noticeably around 60% of free capacity. As mentioned in chapter 6, a school that looses too many students is likely to be closed. Therefore it seems plausible, that a school increases effort sharply when nearing critical values, but also that the teachers give up, when it is clear that the school cannot be saved, as it is likely at a free capacity of 70% or even 80%.

![Figure 8-9: Effect of free capacity on educational productivity](image)

Also interesting is the impact of a variable that measures the predicted outcomes of students, based on the characteristics that schools can observe in new freshmen and while ignoring school productivity. The higher the predicted outcome of the attending freshmen, and thus, the higher the “quality” of the student intake, the lower was the educational productivity.
### Table 8-11: Productivity determination function including predicted scores of student intake

<table>
<thead>
<tr>
<th></th>
<th>educ_prod</th>
</tr>
</thead>
<tbody>
<tr>
<td>competition</td>
<td>1.275*</td>
</tr>
<tr>
<td>(competition)</td>
<td>(0.538)</td>
</tr>
<tr>
<td>compet_sq</td>
<td>-0.168+</td>
</tr>
<tr>
<td>(compet_sq)</td>
<td>(0.0864)</td>
</tr>
<tr>
<td>freecapacity</td>
<td>2.168*</td>
</tr>
<tr>
<td>(freecapacity)</td>
<td>(1.276)</td>
</tr>
<tr>
<td>freecap_sq</td>
<td>-2.381</td>
</tr>
<tr>
<td>(freecap_sq)</td>
<td>(1.860)</td>
</tr>
<tr>
<td>probation</td>
<td>-0.782***</td>
</tr>
<tr>
<td>(probation)</td>
<td>(0.201)</td>
</tr>
<tr>
<td>intake_ass</td>
<td>-0.0636*</td>
</tr>
<tr>
<td>(intake_ass)</td>
<td>(0.0315)</td>
</tr>
<tr>
<td>_cons</td>
<td>7.249</td>
</tr>
<tr>
<td>(_cons)</td>
<td>(4.851)</td>
</tr>
</tbody>
</table>

N: 216
R-sq: 0.137
adj. R-sq: 0.112
F: 5.511

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

An increase in the predicted outcomes of the attending freshmen (intake_ass) by 1 SD reduced the educational productivity by 17% of a SD. This indicates that schools, which are judged based on raw test outcomes of their students, reduce effort if the “quality” of the student intake increases. I will not use this variable in the estimation however. The reason is that this effect is very unlikely to exist, if the quality measure is based on observable educational productivity, as will be the case in the simulation. If educational productivity can be observed directly, then a higher “quality” of attending students will not create the impression of a higher quality of the school. As I need to be capable to compare simulation results for the case where quality measures are based on outcomes to the case where quality measures are based on educational productivity, I have to use the same set of variables for both cases. Therefore, I cannot use the “quality” of attending students in either case.

### 8.8.2. Limit yearly changes in educational productivity

The educational productivity of a school is, apart from the effort exerted by teachers and the principal, based on the average productivity of the teachers, on teaching methods, the curriculum, teaching material and regulations of the school. Most of these determinants of educational productivity cannot be changed at will within a short time. New teaching material has to be found or created, new teaching methods need to be tested, introduced and their application learned by teachers, better teachers can only be recruited when others leave. Therefore, even if the external pressure for educational productivity would rise drastically
within one year, the school might want to increase educational productivity drastically immediately, but the amount by how much it can increase productivity within one year might be limited.

The graph below shows the distribution of changes in productivity from one year to the next that I observed in the CPS for the years in my dataset.

![Distribution of productivity changes from year to year](image)

**Figure 8-10: Distribution of year on year changes in observed educational productivity**

Note that this distribution is based on school-level data, not on student-level data. Therefore, the change in productivity of a school with 20 students in a cohort is represented in the same way as a change in the educational productivity of a school that has a cohort of 400 students. Moreover, these observed changes in productivity contain also changes in unobserved variables and variations in the error term. For small schools, the variation is higher, as averages are based on fewer observations. The changes in educational productivity that I will predict in the simulation do not contain an error term. Also, I do not want the spread of predicted changes in productivity to be widened by estimates for very small schools. Therefore, I will limit the actual change to levels that are slightly below the values that could be justified based on the yearly changes in productivity that were observed in reality.

Of the observed changes in productivity, only about 27% exceeded a change more than +/- 1.0, although the spread of the distribution is likely extended, as explained above. For changes over 2 years, only about 15% exceeded a change of +/- 1.5.
Based on the considerations and findings above, I interpret the predicted educational productivity in the simulation as the intended level of productivity. The actual changes from year to year will be limited to a change of the educational productivity of +/- 1.0 within one year and to +/- 1.5 within two years.

8.8.3. Effects of a directly observable educational productivity

The estimation above is based on a dataset that resulted from a school choice system where educational productivity was not directly observable and parental school choice was strongly driven by average test results. If, as I assume in the simulation, the measure for educational productivity is made available to parents directly instead of the outcome measures used before, the level of educational productivity will have a stronger impact.

If the academic quality measure is based on average test results, educational productivity can only improve the apparent quality of a school indirectly. Moreover, it is possible to create the impression of improvements in academic quality by raising average test results through cheating. Schools can try to avoid academically weak students, can let more weak students drop out of school, focus on students close to meeting state standards etc. If the academic quality measure is based on educational productivity however, most cheating methods do not work anymore. Moreover, the result of an effort-intensive increase in educational productivity is directly observable to school authorities and parents, and will therefore be more rewarding by attracting more applications and avoiding sanctions. This increases the utility that schools gain from increasing effort and should result in higher effort levels.

The fact that increases in educational productivity are more beneficial to schools if educational productivity is directly observable should shift the productivity determination function upward.

8.8.4. Adjusting the productivity determination function to the case of observable productivity

In order to adjust the productivity determination function to the case of directly observable educational productivity, I have to find a level of adjustment that represents the change in a school choice system where observable academic quality is based on outcome to a school choice system where academic quality is based on educational productivity. To my knowledge, there is no study so far that measures this effect. Moreover, as the incentives and dynamics that are generated within a school choice system depend strongly on conditions and
regulations of the individual school system, potential estimates could not be easily transferred to the CPS. Therefore, I have to use an indirect approach that consists of four stages. 1) I survey the literature for findings on the impact of the introduction of incentives for a higher educational productivity, mostly due to school choice, on student outcomes. 2) For each estimate in the literature, I adjust the identified effect on student outcomes, to the best of my knowledge, to the expected change in outcomes when an impact of incentives as the one that results from a choice system as in the CPS is fully reached. 3) I subtract from the average of the adjusted changes in student outcomes the impact of productivity on outcomes that I have already measured for the CPS choice system in its current state. I define the resulting difference in outcomes between the adjusted effects from 2) and real world observations for the CPS as the expected additional impact of choice on outcomes, when academic quality is no longer based on raw test results but on directly observable educational productivity. 4) I alter the productivity decision function of schools, so that the average educational productivity rises to a value that would result in the expected increase of outcomes that resulted from the stages above.

This approach is indirect and involves judgment on my part that I cannot completely base on hard facts. Therefore, I will lean heavily towards conservative estimates in all my considerations, meaning parameter values that lead to a low increase in educational productivity as a result making productivity observable. The result should be interpreted as a lower bound of the effect of making educational productivity observable in the CPS.

8.8.4.(1) Effects of choice on outcomes in the literature

This section contains the first two stages of my approach to altering the productivity determination function.

Lavy (2006) analyzed the effect of the introduction of school choice in Tel Aviv. Each student could choose between 7 schools. Schools could be closed if underutilized. Using a differences-in-differences approach, Lavy found an increase in average scores of 6%. An increase of average outcomes by 6% is in the CPS equivalent to an increase of 77% of a SD in outcomes. This estimate was taken shortly after school choice was introduced. Therefore the full effects of the introduction of school choice had likely not been realized at that time. Moreover, the choice options of students were more limited than in the CPS. Based on these limitations on the impact of school choice, an estimate of an increase in outcomes of 1 SD for
the scenario I intend to simulate for the CPS, namely after many years of school choice and for observable educational productivity, should be conservative. After the introduction of monetary incentives for some teachers in the UK, Atkinson et al. (2004) found an increase in outcomes for the students of the eligible teachers that was equivalent to 73% of a SD. Here, the incentives applied only to individual teachers and the estimate was taken shortly after the introduction of the incentive. Therefore, the increase of outcomes is likely only based on an increase in the effort of the current teachers. All other possibilities to increase productivity, like new teaching methods, a different teaching staff, change of curricula etc. are unlikely to have been employed in such a short time. Also, in the CPS school closure and the loss of employment are possible consequences of bad performances, which should be a stronger incentive than the possibility to earn a bonus. On the other hand, individual consequences are likely to have a stronger impact than group incentives that apply to the entire school, require a collective effort for noticeable improvements and include the possibility of shirking. An increase of 0.75 SD should be a conservative estimate for the case I intend to simulate, where all potential means to improve productivity can be fully employed and where school closures are possible, but with group incentives instead of individual incentives.

Lavy (2004) analyzed the impact of an introduction of monetary incentives for teachers in Israel. He found an increase of the share of students who achieved to earn credits in math by 18% and in English by 17%. I assume that a passing grade in Israel is equivalent to meeting state standards in the CPS. Average scores would have to increase by 1.2% or 1.76 points in the CPS to achieve the same increase in shares as in Israel. This translates into an increase of outcomes by 15% of a SD. The incentives in Israel were limited to neighbourhood schools in disadvantaged areas whose student population should be roughly comparable to that of the CPS, concerning innate ability and socio-economic status. However, in the CPS there are extensive outside options. A large share of students in the CPS attends private schools and, more importantly, the best students can attend selective admission schools. Therefore, the students in the CPS should be comparable to the academically weaker students in the dataset of Lavy. The author differentiated the effect by quartiles based on academic ability, where the first quartile contains the weakest 25% of students. For this quartile, gains were 74% for math and 78% for English. For the second quartile, gains were 50% for math and 30% for English. Gains for the third quartile are not included and the best students showed hardly any gains. The students in the public, non-selective schools of the CPS who are contained in my dataset are likely to be mostly from the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> quartile. Taking this into account, an increase
of outcomes by 0.4 SD for the student population that is comparable to the CPS should be a conservative estimate for the impact of the introduction of choice.

Sandström and Bergström (2002) analyzed the impact of the introduction of state-funded private schools in Sweden on test outcomes. Using an instrumental variable approach for the share of students enrolled in private schools in a municipality, they found that an increase of said share by 10% lead to an increase in average outcomes in public schools of the same municipality equivalent to 22% of a SD. However, this was estimated shortly after the introduction of the funding option, where the full range of measures to improve educational productivity could not have been used. More importantly, the share of students at private schools was still small, with a mean below 5% and a maximum of about 10%. Therefore, the threat to loose many students was rather limited for most public schools. Moreover, public schools had a fixed teacher pool and the curriculum was rather rigid. In the CPS, about 50% of the students do not attend the assigned school and the choice system has been in place long enough so that all potential means for increasing school productivity could be used. Transferred to a situation like in the CPS, with the school choice in place for several years, a much higher threat of loosing students (50% do not attend the assigned school compared to an average below 5%) and threatened jobs, an increase of average outcomes by 1 SD as a result of the introduction of choice should be a conservative estimate.

Hoxby (2000)\textsuperscript{48} estimated the impact of Tiebout choice on the outcomes at public schools. Using an instrumental variable approach based on natural barriers, she found that an increase in the choice index from 0 to 1 was equivalent to an average increase in 8\textsuperscript{th} grade reading by 38\% of a SD. The estimates for 10\textsuperscript{th} grade math were 31\%, for 12\textsuperscript{th} grade reading 58\% and for 12\textsuperscript{th} grade math 27\% of a SD. CPS outcomes are taken in grade 11, with more time for choice to have an effect than for the 8\textsuperscript{th} grade results and less than for the 12\textsuperscript{th} grade results. Taking the average of 8\textsuperscript{th} grade and 12\textsuperscript{th} grade impacts is a conservative approach. As I use average outcomes of several subjects for the CPS, I also average over subjects here. The average of the 4 values above is about 40\% of a SD. The increase from one extreme of the choice index to the other extreme seems to be drastic at first glance. It is, however, equivalent to the difference between the multitude of school districts in Boston and the one large school district in Miami. In the inner city of Chicago, the density of schools is on average probably even higher than the density of school districts in Boston. But more importantly, students in the CPS do not have to move their home in order to get access to another school, as in the Tiebout

\textsuperscript{48} This study of Hoxby was criticized by Rothstein (2005), whose paper was comment upon by Hoxby (2005). This dispute was covered in more detail in a footnote in chapter 4.
choice setting analyzed by Hoxby. Moreover, application procedures and deadlines are coherent in the CPS and information in the one district of the CPS is likely presented more coherently and more easily available than in the multiple school districts in Boston, which should facilitate school choice. Also, the school busing system in one large district is likely more convenient for attending distant schools than the one in an area with a multitude of tiny school districts. Therefore, an estimate of an increase of outcomes by 0.6 SD as a result of the introduction of school choice when transferred to the conditions of the CPS should be conservative.

8.8.4.(2) Expected impact of observable educational productivity on outcomes

The conservative estimates of the impact of the introduction of school choice on outcomes for conditions similar to the CPS show a spread of [0.4, 1] and a mean of 0.75. I will use this mean as a basis for the adjustment of the productivity determination function to the case with observable educational productivity.

Note that this mean of 0.75 is a conservative estimation for several reasons. Firstly, in each case, the adaptation of the estimate to conditions as in the CPS was done leaning towards low values. Secondly, in most of the described studies, the estimates were taken a rather short time after the change in incentives, and in two cases (Lavy 2006, Atkinson et al. 2004) the incentives were targeted at individual teachers. With such a short time, it is unlikely that the majority of schools have already developed and implemented most of the potential cheating measures and many of those measures are much harder to implement for individual teachers, who faced the incentives in two cases, than for the entire school. Therefore, an increase in teacher effort is likely to have driven most of the changes that are captured by the estimates that were described above. Thirdly, educational productivity could not be observed directly in all the school systems described above. If it had been observable, this should have further strengthened the incentives for teachers and schools to increase educational productivity, as opposed to a state where educational productivity only had an indirect effect. In the case that I will simulate, educational productivity will be observable, meaning that it will be more rewarding for schools to increase it. This would further increase the impact of choice on outcomes and educational productivity. That I do not adjust for this fact ensures that the results in my simulation will show a lower bound of potential effects concerning the effects of school choice in the case of observable educational productivity.

The average competitive pressure faced by schools in the CPS was 3.3. The resulting effect of competitive pressure on educational productivity based on the estimation above and compared
to a state without competition (competition=0) is 2.726. This translates into an increase of outcomes by 3.216 which is equivalent to 28.2% of a SD of outcomes. I assume this to be the effect of competitive pressure on outcomes via educational productivity in the CPS system as observed in reality, where the academic quality of schools is based on the meetexceed-outcome measure.

To achieve a total impact of educational productivity on outcomes of 0.75 SD, as conservatively assumed above for the case of observable educational productivity, the productivity determination function has to change in a way, so that an additional increase of outcomes of 0.75-0.282=0.468 SD is reached for the average level of competitive pressure.

I have to admit that the approach that culminates in this number is rather unusual and that this is the part of my calibration that is the easiest to criticize. However, to my knowledge, there is so far no direct estimation of the impact on educational productivity that results from making productivity observable in a school choice system. Moreover, due to the differences in outcomes that can result from seemingly small changes in the conditions or regulations of school choice, findings from any other single school choice system could not be easily transferred to the CPS. Based on the available literature and my dataset, this is the most accurate and sound approach I can come up with. And, as a safeguard against inflated effects, I have throughout this process leaned heavily towards a low increase of the impact of competitive pressure on educational productivity.

8.8.4.(3) Adjusting the productivity determination function

I will now adjust the impact of competitive pressure in the productivity determination function, so that the change results in an average increase of outcomes in the simulation of 0.468 SD, as compared to the case where choice is based on the meetexceed outcome-measure. This calibration is not easily done, as competitive pressure enters the function with the absolute and the squared value. Thus, not only the slope but also the curvature of the impact of competitiveness has to be considered and adapted. The following two facts have additionally to be taken into account. Firstly, in the case with observable educational productivity, the absolute levels of competitive pressure are likely higher. The reason is, that if academic quality measures are based on outcomes, many schools will never be competitive. If the innate academic ability of the assigned students is too low, there is no realistic level of effort that would result in a high share of students who meet state standards. Consequently, such schools will not be considered a competitive threat. In my simulation this is represented by the fact that schools with a quality measure more than 1 SD below average do not increase
the competitive pressure on other schools. In the case with academic quality measures based
on educational productivity however, no school is disqualified as a competitor due to exterior
conditions like an academically weak student intake. This difference alone increases overall
competitive pressure for the average school. Moreover, with observable educational
productivity, the incentive to increase effort is increased. Overall effort levels and therefore
overall productivity will be higher if educational productivity is observable, further driving
competitive pressure up and lifting most schools over the no-threat threshold of 1 SD below
average.

This leads to the second fact that should be considered before adjusting the productivity
determination function: With observable productivity, most cheating methods do not work
anymore. In the case where observable quality is based on test outcomes, schools face
incentives to shift effort from increasing educational productivity to cheating measures. This
is part of the reason for decreasing returns of competitive pressure and the only explanation
why educational productivity should fall with an increase of competitive pressure, as long as a
school is not seen as doomed in any case by its teachers. If educational productivity is
observable, a further increase always eases competitive pressure. The decreasing returns to
scale will however be unaffected. Therefore, the slope of the impact of competitive pressure
should decrease with rising educational pressure, but it would be unrealistic for the slope to
become negative within the levels that are realized in the simulation.

The potential for much higher values of competitive pressure that was described above also
supports the conclusion that the slope of the impact of competitive pressure on productivity
should be decreasing but should not become negative. The reason is as follows: With the
potential for very high levels of competitive pressure, a non-decreasing slope would quickly
lead to unrealistically high levels of educational productivity that increase competitive
pressure further, resulting in a self-reinforcing loop. A decreasing slope however might yield
unrealistically low values of educational productivity, if levels of competitive pressure are
reached that are on the negative part of the function.

Based on the assumption that the slope of the impact of competitive pressure on educational
productivity should be decreasing but not negative, I calibrated the productivity determination
function for the case with observable productivity. This calibration consisted of many trial-
runs of the simulation with incremental changes to the coefficients for the impact of the
absolute value and the square of competitive pressure on productivity until it resulted in an average increase of outcomes close to the 0.468 SD estimated above. The resulting coefficients are 0.74 for the impact of the absolute value and -0.023 for the square of competitive pressure. The graph below shows the impact of competitive pressure on educational productivity for the case of observable productivity:

![Figure 8-11: Effect of competitive pressure on educational productivity](image)

8.9. Imputation of dropout status for the cohort of 2005

The dataset contains information on whether individual students were active in the CPS in the second and third year after entering high-school. For the inactive students, the dataset also contains a leave-code. This leave-code identifies the reason for not staying in the CPS. These reasons can be sorted into three groups: a) the student left the area, b) the student left to another school in the area that is not part of the CPS, like a private school, c) the student dropped out.

Leaving the CPS is not equivalent to dropping out of school. Therefore, in order to identify those students who actually dropped out, I need the leave-codes. But these codes are missing for the entire cohort of 2006 and about half the cohort of 2005. A comparison of the average statistics of the students of the cohort of 2005 who are marked as inactive and miss a leave-code to those students who are marked as inactive and do have a leave-code shows only negligible differences. This fact suggests that the leave-codes had just not been collected and entered into the CPS databases completely when my dataset was generated. Consequently, the share of real drop-outs for the inactive students who still miss a leave-code should be about the same as for the inactive students who have a leave-code.
It is important for the simulations to know which types of students dropped out of school. More precisely, it is important that those students who are marked in the simulation as drop-outs show a very similar profile of characteristics as those students who dropped out in reality. The reason is the following: Students who drop out of school are likely to be academically weak. If a student who is likely to drop out does not drop out, he is likely to have low test results. If the share of drop-outs in the simulation is lower than the share in reality, then the simulation outcomes will include test results for an additional group of students who will have low scores. In this case, the average outcomes for the entire student population would be biased downwards. Then, I could not compare overall outcomes of the simulation to outcomes of the real world school choice system without bias.

For those students who have leave-codes, it is easy to ensure comparability in the two cases by assuming that each student who dropped out in reality will drop out in the simulation. This assumption is not entirely realistic, as it is likely that the dropout share is likely to decrease if overall educational school productivity increases. The alternative would be to predict dropout status in the simulation. But I would not be capable to predict dropout status reliably. The reason is that whether a student drops out is not only determined by his own characteristics but also by school policies which I will discuss below. I could approximate these school policies reasonably well, but only by including school characteristics that are available only for the original data and which I cannot generate in the simulation, like truancy rates. Therefore, I prefer the ability to compare the effects of alternative institutional settings on the official quality measure in the CPS to additional but unreliable insights into effects on dropout shares.

Of the original six cohorts in the dataset provided by the CPS, from 2001 through 2006, I cannot use 2001 because this cohort is missing the ISAT results. If I lost 2005 and 2006 due to missing leave-codes, I would be left with only 3 base-years to feed the simulation. If I could predict the dropout status for those students of the cohort of 2005 who are inactive but miss a leave-code with reasonable precision, I could at least retain the cohort of 2005 and thus have 4 base-years.

As already discussed in 8.1.2, the linear nature of the command “predict” amplifies the effect of characteristics that have a strong effect on outcomes. This effect is exacerbated for the prediction of binary outcomes such as the dropout status. “Predict” can be used for binary estimations, but it does not result in a binary variable but only in probabilities of being a dropout. In order to assign the status “drop-out” to a student, it is therefore necessary to choose a cut-off value for the predicted probability and then to assign the drop-out status to students.
with values above the cut-off. Male students, students from minorities, special education students and students with weak ISAT results are more likely to drop out for example. When using “predict” and a cut-off value to predict drop-out status, the share of students with these characteristics among the dropouts is much higher than in the underlying real world data. The downloadable command “uvis logit” however, replicates the error term of the real world data for the predictions. Moreover, this command generates a dummy variable directly, so that no cut-off values are needed. As a result, the average characteristics of drop-outs based on “uvis” predictions for the cohort 2005 differ only marginally form the average characteristics of those peers who are assigned drop-out status based on leave-codes. Therefore, I impute the drop-out status for those students in 2005 who are missing a leave-code, using “uvis”.

Average individual characteristics of predicted dropouts for cohort of 2005:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>8657</td>
<td>.4088021</td>
<td>.491641</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lunch_status</td>
<td>8657</td>
<td>.88414</td>
<td>.3200754</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lep</td>
<td>8657</td>
<td>.2574795</td>
<td>.437271</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ell</td>
<td>8657</td>
<td>.0220631</td>
<td>.1468972</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_repeat</td>
<td>8657</td>
<td>.6558854</td>
<td>.4751061</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_changer</td>
<td>8638</td>
<td>.2847881</td>
<td>.4513396</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>hs_changer</td>
<td>8657</td>
<td>.3867391</td>
<td>.4870312</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>innate</td>
<td>8657</td>
<td>147.7983</td>
<td>9.948672</td>
<td>120.5</td>
<td>187.5</td>
</tr>
</tbody>
</table>

Table 8-12: Average characteristics of predicted dropouts

Average individual characteristics of dropouts according to CPS data for cohort of 2005 (leave codes were included in dataset):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>6471</td>
<td>.4138464</td>
<td>.4925597</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lunch_status</td>
<td>6471</td>
<td>.8876526</td>
<td>.3158178</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lep</td>
<td>6471</td>
<td>.2560655</td>
<td>.4364922</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ell</td>
<td>6471</td>
<td>.0253438</td>
<td>.1571794</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_repeat</td>
<td>6452</td>
<td>.637923</td>
<td>.480638</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>elem_changer</td>
<td>6471</td>
<td>.2902976</td>
<td>.4539348</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>hs_changer</td>
<td>6471</td>
<td>.3965384</td>
<td>.4892164</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>innate</td>
<td>6471</td>
<td>147.4087</td>
<td>9.782755</td>
<td>120.5</td>
<td>187.5</td>
</tr>
</tbody>
</table>

Table 8-13: Average characteristics of dropouts who are identified in the dataset

The estimations that underlie the prediction by uvis cannot be displayed, but a standard logit estimation with the same variables shows some interesting insights concerning the effects of school characteristics. Table 8-14 shows the “listcoef, std” results that are easier to interpret than direct outcomes of the logit estimation because they give the impact of each variable
measured in standard deviations of the dependent variables\textsuperscript{49} (the interested reader can find the original logit estimation, also differentiation by race, income status or innate ability by running the do-file \textquotedblleft C:\Daten\Stata CPS Data\Graphs and Tables\Table for 8.9. differentiate estimation for dropout.do\textquotedblright):

\textbf{logit (N=73126): Unstandardized and Standardized Estimates}

| Dropout Lost  | b     | \(z\)   | P>|z| | bStdX | bStdY | bStdXY | SDofX |
|--------------|-------|---------|------|--------|--------|--------|--------|
| female       | -0.33471 | -11.907 | 0.000 | -0.1673 | -0.1627 | -0.0813 | 0.4998 |
| lunch_status | -0.18645 | -4.064  | 0.000 | -0.0614 | -0.0906 | -0.0298 | 0.3292 |
| iep          | -0.29883 | -7.222  | 0.000 | -0.1097 | -0.1413 | -0.0533 | 0.3772 |
| ell          | -0.13395 | -1.476  | 0.140 | -0.0186 | -0.0651 | -0.0090 | 0.1386 |
| r_African_Am | -0.37854 | -6.048  | 0.000 | -0.1889 | -0.1840 | -0.0918 | 0.4991 |
| r_Asian_Pac  | -0.64827 | -5.148  | 0.000 | -0.1049 | -0.3151 | -0.0510 | 0.1619 |
| r_Hispanic   | -0.33453 | -5.840  | 0.000 | -0.1611 | -0.1626 | -0.0783 | 0.4817 |
| elem_repeat  | 0.65505  | 22.796  | 0.000 | 0.3266  | 0.3184  | 0.1587  | 0.4986 |
| elem_changer | -0.03798 | -1.233  | 0.218 | -0.0178 | -0.0185 | -0.0087 | 0.4691 |
| hs_changer   | 0.00984  | 0.311   | 0.756 | 0.0049  | 0.0048  | 0.0024  | 0.4993 |
| innate       | -0.03785 | -20.801 | 0.000 | -0.4247 | -0.0184 | -0.2064 | 11.2210 |
| innate_peer  | -0.26431 | -30.943 | 0.000 | -1.3224 | -0.1285 | -0.6427 | 5.0033 |
| educ_prod    | -0.23748 | -12.120 | 0.000 | -0.2209 | -0.1154 | -0.1074 | 0.9304 |
| outcome_att  | 0.20250  | 20.504  | 0.000 | 1.0841  | 0.0984  | 0.5269  | 5.3536 |
| ellatt       | -0.05755 | -2.761  | 0.006 | -0.0582 | -0.0280 | -0.0283 | 1.0111 |
| lowinc_att   | 0.00046  | 0.227   | 0.821 | 0.0060  | 0.0002  | 0.0029  | 13.1929 |
| attend_att   | 0.01030  | 1.983   | 0.047 | 0.0419  | 0.0050  | 0.0204  | 4.0713 |
| mobility_att | -0.01155 | -6.802  | 0.000 | -0.1592 | -0.0056 | -0.0774 | 13.7849 |
| iep_att      | 0.01683  | 3.285   | 0.001 | 0.0892  | 0.0082  | 0.0434  | 5.3011 |
| truancy_att  | -0.01460 | -8.205  | 0.000 | -0.1379 | -0.0071 | -0.0670 | 9.4465 |
| dropout_att  | 0.05322  | 16.661  | 0.000 | 0.3377  | 0.0259  | 0.1641  | 6.3451 |
| graduate_att | -0.00759 | -3.829  | 0.000 | -0.0722 | -0.0037 | -0.0351 | 9.5054 |

\textbf{Table 8-14: Logit estimation of dropout status}

The effects of educational school productivity and average innate ability have the expected sign. A 1 SD increase in educational school productivity (educ Prod) decreases the probability of dropping out by 3% (the measure bStdXY for educ_prod of 10.7\% times the observed SD of the dependent variable for dropout of 0.279). A 1 SD increase in the average innate peer ability (innate_peer) decreases the probability of dropping out by remarkable 17.9\%. But a 1 SD increase in the average PSAE test outcome (outcome_att) at the school in the cohort of the student increases the probability of dropping out by 14.7\%. Differences in average school outcomes in centralized tests are mostly driven by differences in the student composition, captured here by innate_peer as well as other average school characteristics, and educational school quality, captured in this estimation by educ_prod. That a higher average

\textsuperscript{49} The \textquotedblleft listcoef,std\textquotedblright command and how the results are to be interpreted are explained in 8.4.2.
outcome of her peers increases the probability of dropping out for a student, although these and other school-level and individual control variables are included, suggests a suspicious causal relationship.

That better average PSAE scores at the school increase the probability of dropping out for a student although a higher innate ability of her peers strongly decreases the same probability suggests the existence of school policies that increase the probability of dropping out for weak students with increasing scores of their peers. One possible explanation for such an observation would be higher standards. If most students at a school are academically strong, then the required performance level to avoid dropping out might be higher. A second possibility is connected to reverse causality. The scores at the school are higher because more weak students drop out. The official quality measure in the CPS is the share of students who meet or exceed state standards (meetexceed). This share has a strong impact on school reputation and therefore on the number and types of students who apply to a school. Moreover, low meetexceed-shares might get a school on probation status or even lead to closure of the school. Students who dropped out before the test was taken do not count for this measure. As schools are not allowed to directly pick among the applicants and with central tests to determine outcomes, winnowing the weak students by way of drop-out is one of the few remaining possibilities for schools to improve the official quality measure without increasing the educational productivity.

That a higher average test outcome at the school increases the drop-out rate of weak students is an insight that, although interesting, is a by-product of the main line of my research. I cannot examine this phenomenon and the underlying mechanisms more closely within the limits of this dissertation. But this insight confirms my decision not to simulate individual student drop-out, as I cannot replicate the policies of individual schools concerning the treatment of potential drop-outs.

8.10. The estimated model

After incorporation of the findings in 8.1 through 8.9 my baseline model is now complete. Contrary to previous simulations on school choice, this model hardly relies on assumptions. The school choice functions for students, both the function that determines whether a student

---

50 The interested reader can find a little more research on the connection behind an increase in the probability to drop out and the average outcome at the school in “C:\Daten\Stata CPS Data\Explanations & Checks\Explaining dropout.do”.

187
applies to schools at all and the function that determines to which schools she applies, are estimated based on data that was generated by a real world school choice system. Moreover, the data on school characteristics that I use for these estimations are the same data that students and parents had at their disposal when making the school choice decisions. Concerning these decision functions, only the impact of distance and cut-off values to determine whether a student applies to schools at all had to be calibrated. The distribution of distances to the attended school that results from a non-randomized simulation of the calibrated model is very similar to the observed real-world distribution. The chooser shares by race that result from the simulation are also very close to the shares observed in reality, suggesting that the calibration was quite precise.

The productivity determination function for schools was also estimated. I had to develop a measure for competitive pressure in order to estimate this function that depends on the number, distance and relative quality of competing schools. The impact of distance and relative quality are based on observations of real world choice behaviour that approximate the potential threat of competitors to lure students away, based on these two characteristics. To further increase the precision of the model, I differentiated students into groups according to the characteristic that results in the strongest differences in the estimated functions: race. This is important as students differ across types in their choice behaviour and in what determines academic success. Although being far from perfect, the differentiation by at least one characteristic vastly improves the precision of the simulation. As the strongest differences across student groups exist in choice behaviour, this differentiation is especially important concerning results of the simulation about the sorting of students.

The education production function is also estimated based on rather detailed student characteristics including scores in previous tests, peer ability and educational productivity of the school. The estimated function explains three quarters of the variation of test outcomes. The critical part of the model is the adjustment of the productivity determination function to the case of observable educational productivity. Here, I based my approach as far as possible on the data and findings of previous literature. However, it is not possible to base this adjustment only on hard facts so that there is judgment involved on my part, especially when adjusting findings in other school choice systems to conditions that are similar to the CPS. To minimize the risk of inflating the impact of observable educational productivity, I strongly leaned towards lower estimates when making these adjustments. And I did not adjust for the fact that in all the school choice systems that I used, educational productivity was not directly observable so that rewards for a higher effort are lower than in the case that I will simulate for
the CPS. Conservative estimates combined with non-adjustment for the incentive-effects of observability should ensure that the simulation results in a lower bound for the likely effects of making educational productivity observable in the CPS.

8.11. Robustness measures

I deviate from the usual approaches to ensure robustness for this dissertation for four reasons. Firstly, the main findings will be generated by the simulation. The estimations before are only means to this end. Also, the main work that went into this dissertation was the development and programming of the simulation. The preparatory empirical work is merely the calibration of one application of the simulation model. This empirical work is indispensable for reliable outcomes of this application, but it is not the core of this dissertation.

Secondly, most of the variables that were available but that I did not include in the estimations are variables for which there are good reasons not to include them. Truancy rates would, for example, likely have an impact on school choice behaviour. But I do not have student-level data on truancy and could therefore not generate truancy levels for schools in the simulation, when truant students attend other schools.

Thirdly, almost the entire model is based on estimations or calibrated to outcomes in order to be as close as possible to observations made in the real world school choice system. This is a huge step forward as compared to previous school choice simulations that relied mostly on assumptions. Moreover, most of the preparatory estimations are rather precise.

Finally, and most importantly, is the effort that is necessary for calibration. I had to calibrate the impact of distance on choice and the cut-off values for whether a student applies to any school or not. Values in these two calibrations affect each other and I had to calibrate the latter for each base-year of the data and for each race within the base-year. Additionally, I had to calibrate the adjusted productivity determination function in order to result in the estimated increase in average outcomes as compared to the case without observable educational productivity. Between them, these calibrations take hundreds of hours of computing time that require constant control and adjustments. And the whole process would have to be repeated at least partially if values changed due to the common robustness measures, such as the inclusion of another variable or a change in an estimation approach.

I therefore decided to take another approach that increases the robustness of my findings without making additional calibration runs necessary: randomization. Where possible, I randomize the estimated coefficients before using them in the simulation. Where this is not
possible, I use estimations that are based on different base-years and randomly choose which estimation to use in each run.

8.11.1. Randomization of the EPF

I gain robustness concerning the EPF by randomizing the estimated coefficients for all explanatory variables. In a first step, I access the estimation results and read out the estimated coefficient and the SD of the estimate. In the second step, I create a uniform distribution with the spread of +/-1 SD around the estimated value. In the third step, I randomly draw a value from this distribution to be used as the coefficient for this variable in the simulation run. Due to the uniform distribution, each value within the bounds of 1 SD below to 1 SD above the original estimate will be drawn with equal probability. This procedure is done for all explanatory variables separately. Finally, I insert the randomized variable values into the estimation result, using a programming trick that is described in a footnote in 8.1.2. The result is an EPF where the coefficient for each explanatory variable is taken with equal probability from a range of 1 SD above and below the original estimate. This approach creates a wide spread of potential vectors of coefficient values that should cover all combinations of values that are likely to occur in the real EPF with a reasonable probability.

8.11.2. Choice functions

The choice functions have binary outcomes, the student either applies to any school or not and she chose a given school or not. Therefore, I have to estimate the behaviour using a binary approach. The outcomes of binary estimations in STATA cannot be accessed and altered in the way in which I can access and alter the OLS-estimation result of the EPF. Therefore, I have to use another approach for randomization. I estimate the choice functions not only differentiated by race, but also once for each base-year of data. For each simulation run, I randomly and separately draw one base-year estimation result for each choice function. As there are 4 base-years, this randomization results in 16 possible combinations of base-years for the two choice functions.

Concerning the cut-off values above which students decide to apply to any school, I randomize by drawing a value from a uniform distribution from a range that is +/- 0.05 around the calibrated value that applies to the race of the student and the base-year used in this simulation run. These calibrated values are in a range of [0.35, 0.7] with most values around the range [0.4, 0.55]. The share of students who choose their high-school for the case without observable educational productivity, which is the case intended to replicate the real
world CPS, shows an average close to 50% with a SD of 10% and a spread of 20% - 70%. I therefore assume the chosen value for randomization (+/- 0.05) to be sufficiently high. Concerning the impact of distance on choice, I randomize the value $B$ of function (3) by drawing for each simulation run a value from a uniform distribution between 2 and 3, resulting in an average of 2.5, which was the value found in the calibration.
9. The simulation

In this chapter I will present the simulation and the simulation results. First I shortly recapitulate what will be simulated and present how the simulation was technically done in 9.1. In 9.2 I present the two different scenarios for which I will run simulations. Subchapter 9.3 presents and interprets the simulation outcomes. 9.4 concludes with a summary of the findings.

9.1. Simulation approach

I run the simulation based on the estimation and calibration presented in Chapter 8. In order to improve robustness, I randomize most decision functions and calibrated parameters as presented in 8.11. The outcomes of simulation runs include a random element, if only because those of the applicants who are accepted at popular schools are determined randomly. This random student composition then affects school characteristics that affect the school choice of all students in the following turn, altering school choice decisions. Therefore, each simulation run will yield different outcomes. The randomization mentioned above and described in 8.11 intensifies the differences between any two simulation outcomes. One outcome of a single simulation run might therefore differ strongly from the next. Therefore, I base my findings not on the outcome of one single simulation run, but on the average and distribution of the outcomes of hundreds of simulation runs. The average share of active school-choosers, for example, will therefore not be the outcome of one simulation run, but the average of hundreds of simulation outcomes for the share of choosers.

The simulation is run using STATA as described in Chapter 6. Each simulation run simulates the school choice system for 10 consecutive years, where the outcomes of previous years, like the measure for the academic quality of individual schools, are observable in later years and thus affect decisions by the actors. At the end of the simulation of each year, outcomes are gathered and stored in a master-file. Based on these yearly outcomes, I will be able to estimate average effects over a time-span of 10 years and will also be able to see the development of outcome-variables through the years.

For each of these simulation runs, I draw a randomized set of coefficients. Then I run the simulation for two cases for 10 years each, using the same vectors. In the first case, academic quality is measured as it was in the CPS, based on the outcome-measure of the share of
students who meet or exceed state standards. In the second case, academic quality is measured by educational productivity. The result of one simulation run are two simulated 10-year spans that share the exact same set of randomized coefficients and differ only in the nature of the measure for the academic quality of schools. Any findings are based on comparisons between the two cases. That findings are based on the difference between two simulated cases is important for the reliability of these findings. The reason is, that deviations of the simulated model from the real world mechanisms are likely to affect the outcomes in both cases equally. Therefore, basing findings on a comparison of simulated cases mostly insulates these findings from distortions that result from the fact that the simulation model does not represent the real world exactly. An illustrative example for this insulation effect is the following: If predictions of test outcomes in my simulation were too high and I would compare simulated outcomes for the case with observable educational productivity to real world outcomes for the case without observable productivity, then the falsely high predicted test outcomes would seem to be an outcome-improving effect of the observability of productivity. If I instead compare simulated outcomes for the case with and without observable productivity, then the outcomes would be simulated and too high in both cases, negating the inflation of the effect of observability of productivity that is due to simulated outcomes that are too high.

Note, that the simulation does not comprise the entire student population in the CPS. I only simulate behaviour and outcomes for those students who attended the non-selective public schools in the CPS. Students who attended private schools or who attended public schools where admission is based on student ability are not included. Thus, my simulation contains those students from a rather poor inner-city metropolitan area who cannot attend selective admission schools and whose parents cannot, or will not, afford private schools. Those are the students who are trapped in the freely accessible public schools and most likely to benefit from an improved quality.

It is likely that at least some of the students who attend private schools or selective admission schools would attend freely accessible public schools if the productivity there would increase. And it is also likely that some more students would try to leave the freely accessible public schools, should productivity decrease. This, I cannot simulate. Entry into and exit of the public schools is not possible in my simulation. Additionally, changes in the educational productivity are likely to affect the share of dropouts. I do not predict dropout for individual students due to data limitations and because different dropout ratios would make a comparison between cases very difficult. Summing up, I simulate the school choice system
only for those students, who attended the freely accessible schools in the CPS in reality, and whoever dropped out in reality does so in every run and every case of my simulation.

9.2. The two scenarios

I compare the outcomes between the case with observable educational productivity to the case without observable educational productivity for two different scenarios.

9.2.1. Scenario one: passive schools and pure sorting effects

In the first scenario, schools are inactive, meaning that their educational productivity does not react to external pressure. In this scenario, the educational productivity of a school in each year is fixed at the value that was observed in the CPS for the base-year that provides the data for this simulated year.

For the case without observable educational productivity, I run the simulation as calibrated, but with educational productivity fixed, instead of determined by the estimated function. For the case with observable educational productivity, I run the simulation as calibrated but replace the measure for academic quality that was used in the CPS, the share of students who met or exceeded state standards, with a measure based on educational productivity. This measure was normalized to a mean of 0 and a SD of 1, as was the outcome-based measure before I estimated and calibrated the model. Due to this normalization, I can replace the original measure with another measure with the same mean and SD, so that differences in the outcomes of the model between the two cases are not affected by possible differences in the mean or variance of the measure for school quality. By simply substituting the original measure for academic quality with a measure based on educational productivity, I assume that parents react to the new measure in the same way as they reacted to the old measure. I assume that the new measure replaces the old one, meaning that the old measure is no longer published. Parents often have scant understanding of the mechanisms used to compute quality measures for schools. They base their assessment of the academic quality of schools on the official measures that are easily accessible, on school report cards or in league tables or other rankings in newspapers or the internet. They are therefore likely to react in the same way to a new measure, if this measure shows familiar average values and familiar differences between schools, especially when this measure for academic quality is the only one that is published51.

---

51 There are two possible concerns with assuming that the new quality measure will be used by the parents in exactly the same way. First, parents might not be interested in choosing a school with a high educational
As schools cannot adjust their educational productivity in this scenario, any observed effects result from changed school choice actions of the parents and students. Even effects on average outcomes that result from the observability of educational productivity are therefore due to sorting of students into the schools with the highest productivity.

9.2.2. Scenario two: educational productivity determined by external pressure

In the second scenario, schools are active, meaning that they react to external pressure by changing their effort and thus their level of educational productivity. The productivity determination function is active in this scenario and determines the educational productivity of schools.

For the case without observable educational productivity, I run the complete simulation as calibrated in chapter 8. This simulation contains the information about schools as parents could observe it when making school choices and represents the real world choice system of the CPS as closely as possible. For the case with observable educational productivity, I again replace the meetexceed measure for academic quality that was used in the CPS with a quality measure based on educational productivity. As in the scenario before, this measure is normalized to the same mean and SD as the meetexceed-measure that was used to estimate the decision functions. And, as in the first scenario, I assume that the new measure for academic quality replaces the old one, meaning that the old measure is no longer observable.

In this scenario, both sorting effects and the effects of changed incentives for schools have an effect on the school choice system. This scenario yields the most realistic estimate of the effects of introducing academic quality measures that are based on educational productivity in the CPS. Note however, that due to the conservative adjustment of the productivity

productivity when choosing high-outcome schools. Instead, they might be interested in the “right” peers. However, in the CPS school report cards, a lot of information on the peer composition is included. The racial composition, low income shares and IEP shares are published and enter the school choice equations of my simulation. Additional information, that I could not incorporate into the simulation was given with dropout rates, truancy rates, attendance rates and graduation rates. With this information, parents could infer detailed information on the potential peers of their children, without having to use the average outcome as a proxy. The second concern relates to those parents who care about and understand the mathematics behind the old CPS outcome measure. Value-added measures are harder to understand (Wilson et al. 2004). Therefore, a reporting format should be taken that makes it relatively easy to interpret these indicators (Meyer 1997). By using a quality measure that is standardized to the mean and spread of the old quality measure, I ensure that it is easy to interpret for those who are used to the old meetexceed measure. Understandability can be increased by an information effort. The CPS informed parents about a new value-added measure for elementary schools, among other things with a power point presentation (CPS Value Added). As information about peer characteristics was available before through direct variables, it is unlikely that an increase in the precision with which the academic quality measure shows educational productivity would reduce the reliance of parents on this measure when choosing a school.
determination function, the findings will likely represent a lower bound for the effects that can be expected to result from the introduction of productivity-based quality measures.

9.3. Simulation runs and outcomes

In this subsection I show the outcomes of the simulation runs of the two scenarios. Note, that all outcomes that I present here, are of two types. I either present distributions and/or averages of all the turns and all the runs for each case and scenario. Or I present averages by case and scenario differentiated by the turn of the simulation from which they resulted. The first type of outcomes shows the average outcome over the entire time of 10 simulated years. The second type of outcomes presents the effect by year, so that developments during the simulation can be observed. Both are based on the average of hundreds of simulation runs.

Also, remember that all insights are based on comparisons between two cases, not on raw outcomes. Most insights are based on comparisons between the simulation-outcomes of the two cases. These comparisons show the effect of replacing the meetexceed measure for quality that was used in the CPS with a quality measure based on educational productivity. The second type of comparisons is between a simulated outcome and a hypothetical situation where all students have to stay at the assigned school. This means, that the average innate peer ability is frozen at the level that would occur without choice. The educational productivity of schools is also frozen at the level that was observed in reality for the corresponding base-year. The comparison between these outcomes and simulated outcomes shows the impact of choice, including the effects of choices in previous simulated turns. However, even this second type of comparison does not show the full impact of an introduction of school choice into a school system without school choice. The reason is that the existing school choice and the resulting external pressure in the CPS induced adjustments in educational productivity by schools, even though educational productivity was not observable. With the approach I use it is not possible to find the impact of an introduction of school choice without heavy reliance on assumptions. The reason is that choice has to be observable in order to estimate the reaction functions of both schools and students. If this is the case, then there is no data on the same school system without choice, at least no recent data. In order to be able to estimate the total impact of an introduction of school choice, it is necessary to repeat the approach of this dissertation for several other school choice systems. I will elaborate on this possibility in chapter 10.
However, the findings that I present for the second scenario are likely not too far from the total impact of an introduction of school choice. I will elaborate on why this is the case after presenting the findings of the first scenario.

9.3.1. Simulation results for scenario one: passive schools

Now I present the effects of introducing an academic quality measure for schools that is based on educational productivity in the CPS. Here, the educational productivity is fixed, so that effects can only result from changes in school choice decisions and the resulting sorting of students into schools. In presenting these results I will focus mostly on overall effects, averaged over students from different races and all levels of innate ability. The interested reader can find differentiated results by running the do-file: “C:\Daten\Stata CPS Data\Graphs and Tables\For 9.3.1. creating graphs final”. This file contains command lines that show most of the results for gains of innate ability, scores and productivity for active choosers and non-choosers, differentiated by race and innate ability.

9.3.1.(1) Effects on Productivity

The graph below presents the distribution of the gain in productivity for the average attended school, as compared to the case where every student has to attend the assigned school. I compute this number by subtracting the average educational productivity of the assigned school from the average educational productivity of the attended school. The resulting number is therefore the average gain in productivity that results from choice in one turn of one simulation run. This number is represented as one of the observations underlying the graph presented below. The translucent (black) bars represent the distribution of outcomes for the case where academic quality was measured using the meetexceed quality measure. The x-axis measures the percentage of observations that falls within the quantile of observations that is covered by the bin underlying the corresponding bar. The filled (red) bars represent the corresponding distribution for the case where the measure for academic quality was based on educational productivity. The graph is to be interpreted as follows: a filled (red) bar for a productivity gain of X that reaches the height of 2 percent means, that in 2 percent of all the simulated turns for the case with observable productivity, the average gain in productivity fell in the bin that contains the value X. The graph also includes lines at the mean of the gain in productivity for both cases.
Figure 9-1: Scenario 1: Distribution of average gains in productivity as compared to a scenario without choice

The distribution in the graph shows that for the case with quality measures based on educational productivity, called PROD-case from now on, the average gain in educational productivity was 0.132 points. This means that, in the average over all turns of all simulations, the average attended school had a productivity that was by 0.132 points higher than if all students had attended the assigned school. The distribution of values for individual turns is wide, but mostly above 0 and clearly to the right of the distribution for the case where quality measures were based on the meetexceed share, called OUT-case from now on. In the OUT-case, the average over all observations is -0.068, meaning that in this case, students sorted themselves into schools that have on average a slightly lower educational productivity than the assigned schools. Here, most of the observations are below 0.

This difference between the cases becomes even clearer, if I only analyze those students who did not attend the assigned school. The distributions of gains in educational productivity for these school-changers are presented in the graph below.
In the PROD-case, the average gain in educational productivity of school-changers, on average over all turns in all simulation runs, is 0.378 while in the OUT-case, school-changers an average lose 0.119 points of productivity. This finding fits well with two facts presented above. Firstly, Cullen/Jacob (2007) found that school-changers in the CPS strongly followed high test outcomes when choosing schools and that among those schools that attracted applications, those with a higher productivity were less popular. Secondly, a higher innate ability of attending students leads to a lower educational productivity. Based on these observations, the following explanation for the difference in productivity gains between the two cases seems likely: In the OUT-case, parents and students cannot observe educational productivity. In search of a school with a high academic quality, they rely on the observable measure when choosing schools. This measure is mostly driven by the student composition. Consequently, school-changers apply to schools with a high average innate ability. These schools do not necessarily have a high educational productivity. Moreover, when schools are judged by test outcomes and a school has a student intake that guarantees high average test scores, schools ease their effort resulting in a lower educational productivity (see 8.8.1). As school-changers tend to have a high innate ability and they apply to schools which show high test scores that are mainly due to a high average innate ability, these...
students attend schools that have high test scores but a rather low educational productivity. The result is that these students are rewarded for their school change on average with a loss in educational productivity. The graph below shows the development of average gains in productivity over time, aggregated by the 10 turns of each simulation run and presented for all students, not only for school-changers. As can be seen from the trajectory, this development is rather flat. In the PROD-case, when educational productivity measures are made publicly available, school-changers can immediately identify the schools with a high productivity. They attend these schools, resulting in a higher average educational productivity of the attended school.

Note that the saw-tooth pattern of outcomes over the turns is an artefact of the limited number of base-years at my disposal and the fact that I always have to start with the same base-year. The explanation in more detail: my student-level data are organized by cohort. I have full data for only 4 cohorts from the base-years 2002-2005. As the choices of the actors are based on aggregate school characteristics from previous years, I need observations for a few years before the year that I use as the base-year for the first cohort of students to choose schools. The educational productivity for a cohort can only be computed after this cohort has taken the PSAE test in 11th grade, and is then observable when the next cohort chooses schools.
Therefore I need three years of observable data before the starting cohort. The productivity of my first cohort, 2002, is only visible when the cohort of 2005 makes school choice decisions. Therefore, I can only start with one base-year: 2005. Then, I use the base-years again in the order in which they occurred in reality, starting with the data from 2002 after using the data from 2005 and filling in the aggregate school characteristics with values that I compute as an outcome of the simulated turns. A randomization of the order of base-years is also not possible\(^{52}\). Consequently, each simulation run will have to start with the base-year 2005, followed by 2002, then 2003, 2004, 2005, again 2002 and so on.

The base-years differ. The total number of students varies, as do school characteristics including educational productivity. This leads to outcomes that differ with the base-year that provided the data, resulting in a distinct and regular saw-tooth pattern for some outcomes when observed over time.

**9.3.1.(2) Effects on test outcomes and average innate ability of peers**

With the gains in productivity presented above, it is no surprise that average gains in scores for the entire student population are on average slightly negative for the OUT-case at -0.055 or 0.5\% of a SD and positive for the PROD-case at 0.167 or 1.5\% of a SD. Gains for both cases are outcomes of the corresponding case as compared to predicted outcomes if all students had to attend the assigned schools. These gains are not large. The difference between the two cases is only about 2\% of a SD of outcomes. However, this is only a result of the sorting of students into the most productive schools, with no impact of the change in quality measures on the educational productivity of schools.

\(^{52}\) The reason is that I also have to use school characteristics from the case without school choice to compute several outcomes. And I have to keep track of which student attends which school to compute some school characteristics. The risk of making programming mistakes and getting school characteristics from the wrong years would be too high if I randomized the order of the base-years. Moreover, if I randomly chose the next base-year, it will happen that the same student is attending high-school several times at once. As I store several variables with the student data, I would have to multiply each observation. Also, I would have to repeatedly switch students from an active status to an inactive status or, alternatively, delete them from the dataset and get them back in. Each possibility would be messy, take up computing time and greatly increase the risk of making programming mistakes. With the fixed order of base-years, a student has left high-school before her base-year is used again, which enables me to evade all the problems described above.
Figure 9-4: Distribution of average gains in test scores as compared to no-choice scenario

The distribution of average score gains for individual simulated turns is mostly to the right of 0 for the PROD-case and more than half of the distribution of the OUT-case is to the left of 0.
The development over the simulated years again shows a saw-tooth pattern but no clear trend, indicating that any effects on outcomes are immediately realized when students sort into the most productive schools as a reaction to observable educational productivity.

The graph below shows the distribution of score gains separated for those students who actively chose schools and for those who stayed at the assigned school.
The gains for students who actively choose schools are higher than the average gains for all students. In the PROD-case, the average gain over all turns for the average school-changer is
0.761 points or 6.7% of a SD of test outcomes. Even for the OUT-case, the average score gain for school-changers is 0.268 points or 2.3% of a SD of outcomes. With fixed productivity levels, the score gains of school-changers for both cases in this scenario can result only from sorting into the limited places at the most productive schools and from a higher innate peer ability. In the OUT-case, the high-productivity schools cannot be identified and are only chosen by chance, leaving only a change in average peer ability to systematically affect test scores. The resulting change in scores in the OUT-case is rather small, which might explain why previous studies hardly found an effect of active school choice on the average academic outcomes of students. Note however, that this scenario is meant to present the isolated effects of the sorting of students into schools. The reactions of schools to competitive pressure that I estimated for the real CPS are excluded in this scenario by fixing the educational productivity of schools at the levels that were observed in the CPS. The activation of the productivity determination function in the next scenario might lead to different results for the OUT-case. Therefore, the exact values in this scenario should not be given too much weight when explaining observations based on real world data.

In the OUT-case, school choices are strongly driven by average test results that in turn are strongly determined by the average innate ability of students. Therefore, school-changers in the OUT-case change to schools with a high average innate ability. As high-ability students are more likely to change schools, those who stay at the assigned school loose the peers with the highest ability. As a result, the students who do not change school show lower scores than in the hypothetical scenario where all students have to stay at the assigned school. The average non-changer student looses in the OUT-case 0.335 points or 3.1% of a SD of outcomes. The loss in the PROD-case is 0.189 or 1.7% of a SD.

The graph below shows the change in average innate ability that results from choice in the two cases, separated for school-changers and non-changers. Remember that these numbers result from a comparison of simulated outcomes to the hypothetical scenario where all students have to attend the assigned school. These numbers do not show the effect of actively choosing a school rather than attending the assigned school for an individual student in one turn.
Figure 9-8: S1: Gains in average peer ability, distribution for choosers

Figure 9-9: S1: Gains in average peer ability, distribution for non-choosers
In the PROD-case, the average school-changer gains 2.251 points or 45% of a SD of innate peer ability, in the OUT-case, the gain is 2.793 or 56% of a SD. The corresponding losses of the drain of high-ability students on the peer ability of non-changers is 1.453 or 29% of a SD of innate peer ability in the PROD-case and 2.509 points or 51% of a SD in the OUT-case. Summing up, the score gains of school-changers in the OUT-case are rather limited, mostly based on a higher average peer ability and come at the cost of a lower average peer ability, and thus score losses, for non-changers.

**9.3.1.(3) Share and characteristics of school-changers**

The next two graphs show the distribution of the share of school-changers and the average innate ability of a school-changer.

![Distribution of share of choosers](image-url)
In the OUT-case, the share of students who does not attend the assigned school, averaged over the shares of all turns in all simulation runs, is 47.42% and thus very close to the share observed in reality of 47.87%. The share of school-changers in the PROD-case is lower at 39.41%. The average innate ability of the school-changer is higher in the PROD-case than in the OUT-case. If the costs of changing schools, like a longer commute and being separated from former classmates, were taken into account, these would be lower on average in the PROD-case, further increasing the advantage in average scores of the PROD-case.

The reason for the lower share of school-changers in the PROD-case is likely that many schools in areas with academically weak students are better than they seem to be according to the meexceed measure. In the CPS, two thirds of the school-changers are African Americans. The reason for this high share is mostly, that the assigned schools of African American students are in areas with academically weak students. As a result, the assigned schools score badly on the meexceed quality measure. The average quality of the assigned school for African American students is, according to this measure, -0.452, as compared to 0.423 for White, 0.113 for Asian and -0.081 for Hispanic students. Thus, according to the

---

53 Of the African Americans, 59% are school-changers. The corresponding share for White students is 39%, for Asian students 55%, for Hispanic students 33%. Add to that, that 54% of all students in the CPS are African American, 36% are Hispanic, only 7.5% are White and 2.5% are Asian. Therefore, of the 39,419 school-changers in the CPS, 26,352 are African American.
measure for academic quality used in the OUT-case, most African American students have to change schools in order to get access to a school with a decent quality. This quality measure is deceiving however. When looking at the educational productivity of the assigned schools, then African American students are assigned to schools with an academic quality slightly above the average. Therefore, making educational productivity observable unmasks, that the group of students who provides that largest share of school-changers is, in reality, not assigned to bad schools but rather to schools with an acceptable academic quality. The share of school-changers drops strongly, by about 13.5% for African Americans when comparing the PROD-case to the OUT-case. The corresponding shares for the other students remain mostly unchanged.

Regarding the innate ability of students confirms the findings above. While the share of school-changers for high-ability students hardly changes when comparing the PROD-case to the OUT-case, the share for low-ability students decreases from 17.7% to 11% and for medium-ability students from 50% to 39.6%. To sum up: when educational productivity becomes observable, disadvantaged students can observe that their assigned school is often not as bad as it seems to be based on average test results. Moreover, disadvantaged students can observe that many schools with high average scores only seem to have a high academic quality and that changing to that school would often mean a decrease in academic quality. As a consequence, the school-changer share of disadvantaged students drops, decreasing the overall share of school-changers. The fact that the average innate ability of school-changers is higher in the PROD-case results from the fact that the share of high-ability students hardly changes while less low-ability students change schools.

9.3.1.(4) Sorting by ability

One of the concerns about school choice is, that it might lead to increased sorting of students by ability, race/ethnic background or family income. These concerns are justified, as students who are socially advantaged and/or have a higher innate ability are more likely to choose schools actively. Therefore, they might accumulate in a few choice schools, leaving the academically and socially weakest students behind. The following graphs present the distribution of sorting outcomes for all turns of all simulated runs, separated by case. The measure I use is the share of students of this group that would have to change schools, so that all schools would have the same share of this type of student. A measure of 0.3 for high-ability students would mean that 30% of the high-ability students would have to change schools in order to achieve an equal share of high-ability students at all schools. A reminder:
high-ability students are those students whose 8th grade ISAT test scores fall into the quantile containing the 20% of students with the highest test scores. Low-ability students are in the quantile that contains the 20% of students with the lowest test scores.

In the graphs on sorting, I added a third (blue) line. This line represents the sorting measure for the hypothetical case where school choice is not allowed and all students have to attend the assigned school.

Figure 9-12: Distribution of sorting by innate ability for high ability students
As can be seen from the graph, sorting is higher in both cases as in the case without school choice. Sorting is however slightly lower in the PROD-case than in the OUT-case for high-innate students (27.0 vs. 28.7) and clearly lower for low-ability students (25.4 vs. 32.6).

9.3.1.5 Sorting by race and income
Concerning sorting by race, both the PROD-case and the OUT-case show sorting by race that hardly differs from the distribution according to the assigned schools. Note, that sorting by race for these two groups that are usually academically disadvantaged is quite high, around 70% for African American and around 65% for Hispanic students, reflecting the fact that students of both groups tend to cluster strongly in the CPS, attending schools where their own group has a vast majority.
The graph below shows the distribution of sorting by race for White students and sorting by income, based on whether students received lunch for free or at reduced prices.

Sorting by race for White students is in both cases of school choice higher than for the hypothetical case without school choice. Sorting by income-level hardly differs across the three cases. Only the variation is higher for the PROD-case.
9.3.1.(6) Capacity utilization of schools

The following graphs show the average capacity utilization of schools over the 10 years of each simulation run. The quality measures in both cases can only be computed, when the corresponding cohort has taken the PSAE test in grade 11. Therefore, school choice decisions have to be made based on a quality measure that represents the quality as it applied to the students of three cohorts earlier. Meanwhile, the quality of schools might have changed. Two insights are interesting concerning the capacity usage. Firstly: what types of schools did students want to attend? This question can be answered using the quality measures that students could observe when making their school choice. Secondly: what types of schools did students attend? This question can be answered by using the quality of measures that applied to the cohort of the students and are only visible more than two years after the choice decision was made.

The first set of graphs shows the share of free capacity at the schools that had an educational productivity at or above the average. The left graph is based on the quality measures of the schools that were observable when taking the school choice. The right graph is based on the quality measure of the attended schools that applied to the cohort of the student.

Figure 9-16: S1: Average share of free places at more productive schools by turn, based on expected and experienced level

Based on measures that were observable when the choice-decision was made, in the PROD-case, the free capacity of those schools with educational productivity at or above the average is at 10.2%. In the OUT-case, around 37.1% of the capacity at the schools with a better educational productivity is free. Based on the measures that actually applied to the cohort of
the choosers, the free capacity at the schools with higher productivity is at 23.3% in the PROD-case and at 37.2% in the OUT-case. The fact that an additional 14% of the capacity at the more productive schools is filled in the PROD-case accounts for the differences in average educational productivity of the attended school and in test scores between the two cases that were described above.

The second set of graphs shows the share of free capacity at the schools that had a meetexceed share at or above average. The left graph is based on quality measures observable when school choice decisions were taken, the right graph is based on the quality measures that applied to the cohort.

![Figure 9-17: S1: Average share of free places at high-scoring schools by turn, based on expected and experienced level](image)

The share of free capacity at schools with a meetexceed share at or above the average, based on values observable when taking the school choice decision, is around 11.9% for the OUT-case and around 33.7% for the PROD-case. The share of free capacity at schools with a high meetexceed share based on the quality measures that apply to the cohort is around 12% for the OUT-case and around 18.1% for the PROD-case.

### 9.3.1.(7) Implications of findings of scenario 1 for scenario 2

As can be seen from the findings of Chapter 8 and the results of the first scenario, school choice as it happens currently in the CPS, with academic quality measures based on raw outcomes, does not lead to students striving to get into the most productive schools. This has
two implications. Firstly, the possibility of choice does not lead to students going on average to more productive schools, cancelling positive effects of choice on outcomes that result from sorting. Secondly, as can be seen from the shares of free capacity, having a high educational productivity does not lead to a significant advantage for schools, meaning that there are at most weak incentives to put effort into an increase of educational productivity. Therefore, the effects of choice on educational productivity were likely rather weak in the real CPS. The outcomes in the OUT-case should therefore not change strongly when productivity is determined in the next scenario by the productivity function, as this function was estimated with CPS data. In the PROD-case, the difference between the scenarios should be stronger, as the observability of educational productivity increases the incentive effects. The difference between the two cases should be stronger in the next scenario, as both the sorting effect and the incentive effect of observable educational productivity will have an impact. Moreover, these two effects might reinforce each other. An example for this mutual reinforcement is the following: If the incentive effect leads to higher levels of educational productivity in some schools, these attract more students, increasing the impact of the sorting effect. If some schools attract more students, the free capacity at other schools increases, inducing these schools to increase effort. These reinforcing effects should lead to a trajectory of the total effects of choice across time that is not flat as in the first scenario.

9.3.2. Simulation results for scenario two: educational productivity determined by external pressure

In the second scenario I present again the effects of introducing in the CPS an academic quality measure for schools that is based on educational productivity. Contrary to the first scenario, the educational productivity is no longer fixed, but determined by the productivity determination function that was estimated in chapter 8.8 and thus by competitive pressure, free capacity and the probation status of individual schools. In the OUT-case, the productivity determination function is as estimated. For the PROD-case, the function is adapted to the incentive effects of the observability of educational productivity as presented in chapter 8.8.4.(3). In presenting these results I will focus mostly on overall effects. The interested reader can find differentiated results by running the do-file: “C:\Daten\Stata CPS Data\Graphs and Tables\For 9.3.2. creating graphs final”. This file contains command lines that show most of the results for gains of innate ability, scores and productivity for active choosers and non-choosers, differentiated by race and innate ability.
9.3.2.(1) Effects on Productivity

I begin by presenting the effects on the average educational productivity of the attended school. It is important to note here, that I do not gain this result by comparing the simulated productivity value of the attended school for the current turn to the simulated value of the assigned school of the same turn. Instead, I compare it to the productivity value observed for this school in reality in the corresponding base year. There are two reasons for this approach. The first reason is to ensure comparability of gains to the first scenario. There, productivity for the assigned schools is fixed to the values observed in the corresponding base years. If I fix productivity of the assigned schools to these values as well in the second scenario, then I can compare the size of gains between the two scenarios.

The second reason for using base-year values for the productivity of the assigned schools is to identify the overall effects of the observability of educational productivity in a dynamic simulation. In the simulation of this second scenario, the educational productivity of schools is determined by choice and competition and thus changes from year to year. If I updated the productivity of the assigned schools by using the simulated values for the current turn instead of the base-year values, the gain measure would only identify the additional effect of prolonging the already existing observability of productivity by another year. By using the base-year values, I can identify gains as compared to a situation where choice was never introduced, and thus measure the overall dynamic effects of school choice over several turns. The interested reader can find the actual average gains for individual students that result from the choice of actively choosing a school instead of staying at the assigned school in one turn in the NCTT values of the outcome-file. NCTT stands for “No Choice This Turn” and compares results for individual choosers for the case where they actually attend the school that they choose according to the simulation to a hypothetical case where only this individual student stayed at the assigned school, while all other actors behaved as simulated. NCTT results therefore show the difference between choice and no choice for the individual student in the simulated turn. NCTT results exist for productivity (out_gain_prod_sw_nctt), scores (out_gain_sc_), average innate ability of peers (out_gain_in_) and the meetexceed share (out_gain_psae_). All those NCTT gains are also included in the output-file differentiated by race. In the first scenario, the productivity values for the assigned schools remain fixed throughout the simulation at those values that were observed in the CPS. In the first scenario, the NCTT gains regarding productivity are therefore identical to the gains presented in

54 These results can be accessed by running the do-file: “C:\Daten\Stata CPS Data\Graphs and Tables\For 9.3.2. creating graphs final”.

216
9.3.1.(1). The NCTT gains for productivity in the second scenario differ from the ones presented above. They are -0.031 for the OUT-case and 0.69 for the PROD-case, meaning that the average chooser attends on average a school with a slightly lower educational productivity than her assigned school, when choice is based on the meetexceed-measure.

The following graph shows the distribution of gains in the average productivity of the attended school as compared to a hypothetical situation where all students had to attend the assigned school. As for the first scenario, the X-axis measures the percentage of outcomes for individual simulated turns in which the gain in productivity (averaged over all students) falls into the bin underlying the corresponding bar.

![Average gains in educational productivity](image)

*Figure 9-18: Scenario 2: Distribution of average gains in productivity as compared to a scenario without choice*

For the OUT-case, most of the distribution of outcomes for individual simulated turns is below 0, while all observations for the PROD-case are above 0. The average for the PROD-case is 1.429, meaning that the average gain in productivity, averaged over all simulated turns and as compared to a situation where all students have to attend the assigned school, is 1.429. In the OUT-case, the average loss in productivity is -0.394. As educational productivity was
standardized to mean 0 and a SD of 1, these values correspond to changes as measured in SD of productivity.

In the first scenario, only the sorting of students into the more productive schools can affect the gains in productivity. In this second scenario there is additionally the incentive effect that induces schools to increase educational productivity in response to competitive pressure, free capacity and probationary status. Compared to the first scenario, the productivity-gain in the PROD-case increases by about 1.3 points, the loss in the OUT-case increases by about 0.3 points. The difference for gains in productivity for the average attended school between the two cases increases from 0.2 SD to 1.8 SD. According to these findings, the rather limited difference between the two cases in the first scenario is strongly increased, when the incentive effect and potential cross-effects are added to the effect of sorting into the more productive schools. The clustering of observations for the PROD-case is a result of a gradual increase in productivity over time. I will elaborate below on this clustering.

The following graph shows the average productivity-gains for those students who did not attend the assigned school.

![Gains in educational productivity for choosers](image)

**Figure 9-19: S2: Distribution of average gains in educational productivity for choosers**

In the PROD-case, the average over all turns for the gain in productivity for those students who did not attend the assigned school and as compared to the hypothetical scenario without
choice is 1.459. The corresponding number for the OUT-case is -0.459. As in the first scenario, the effect for active school choosers is more pronounced than for the average student. In the PROD-case, the gain is marginally higher, in the OUT-case, the loss is a little higher. The difference in productivity gains for choosers increases from 0.5 SD in the first scenario to 1.9 SD in this second scenario.

The following graph presents the gains in productivity by turns of the simulation:

![Educational productivity gain by year](image)

**Figure 9-20: S2: Average gain in productivity by year as compared to scenario without choice**

The values start at 0.179 for the PROD-case and -0.037 for the OUT-case, close to the average observations for the first scenario (0.132 and -0.068).

The productivity determination function is driven by competitive pressure, free capacity and probationary status. Probationary status is fixed to the values observed in the real CPS. Competitive pressure is a function of the number, distance and academic quality of other schools. The number and distance of other schools are fixed, as the closure and foundation of schools is not endogenous in my model. The academic quality is in both the PROD-case and the OUT-case computed based on 11th grade PSAE results. Only in the 4th turn are results of the students who entered high-school in the 1st turn visible. In the first three turns, the values for competitive pressure that are observable to schools are therefore determined by outcomes of students who had entered high-school before the beginning of the simulation. I assume that
any changes to productivity that occur in response to competition only affect the students who enter high-school in, or after, the year when the changes are introduced. Therefore, the levels of competitive pressure in the first three turns are unaffected by simulation results. Only in the 4th turn can the effects of simulated decisions in previous turns affect the value for competitive pressure. As probationary status is fixed and competitive pressure is unaffected by choice decisions in the first three turns, the only channel for choice to affect productivity in these turns is through changes in the free capacity of schools.

Starting with the 4th turn, the dynamic impact of the incentive effect can start to work, as the productivity determination function is now fed by values that were determined by the productivity determination function and school choice decisions in previous simulated turns. In other words: changes in the educational productivity of competing schools, that were a reaction to choice decisions under the altered circumstances but were invisible so far as the affected students had not yet taken the PSAE test, now become visible and provoke reactions. In the PROD-case, this reaction is for most schools to increase productivity. The desired increases in productivity cannot be affected at will within one year. As presented in chapter 8.8.2., I limited changes within one year to +/-1 and within two years to +/-1.5 points.

That the observations for productivity gains form clusters, based on the turns in which the observations occurred, can be explained by three facts. The first fact is the limit to yearly changes that effectively binds actual changes for a non-negligible share of schools. If most schools desire, as a result of changed incentives, to alter their educational productivity strongly, the limitation of yearly changes is effective and restricts these changes for all schools. The second fact is that several of the determinants for educational productivity (probationary status, number and distance of competing schools) are not endogenous but fixed to the values that were observed in reality in the CPS. These variables are therefore determined by the base-years that always underlie the same turns of the simulation, as every simulation run has to start with the base-year 2005. The third fact is, that the endogenously determined variable with the most impact on educational productivity, the educational quality of competing schools, becomes observable only 3 turns after the value is chosen by the competitors. Consequently, the observable values in the first 3 simulated turns are unaffected by simulated decisions and thus identical in each simulation run, leading to similar responses in turns 4 to 6, when decisions are based on observations in the first 3 turns. The reactions of schools in turns 7 to 9 then are affected by the responses in turns 4 to 6 to decisions in turns 1 to 3.
The average growth of productivity gains is rather weak in the first turns, rather strong in the turns 3 to 7 and again weaker after turn 7. This pattern is due to the fact described above, that reactions to changed conditions, such as a changed school quality measure, become observable only with a delay of several years. This delay might also help to explain, why studies that were undertaken only a few years after the introduction of school choice or other changes in incentives that aimed at a higher educational productivity often did find no or only small effects.

9.3.2.(2) Effects on test outcomes and average innate ability of peers

In the second scenario, average score gains for all students in the OUT-case are again negative at an average over all simulated turns of -0.513 points or 4.5% of a SD of outcomes. Average score gains for the PROD-case are 1.616 points or 14.2% of a SD of outcomes. The difference between the two cases is 2.129 points or 18.7% of a SD of outcomes, up from 2% in the first scenario, where only the sorting of students into the most productive schools could have an impact on the productivity of the attended school.

![Figure 9-21: S2: Distribution of average gains in test scores as compared to no-choice scenario](image)

The distribution of score gains for the OUT-case is mostly below 0, the distribution for score gains in the PROD-case is mostly above 0 and the distributions for the two cases are more clearly separated than in the first scenario.
As for the gains in productivity, the values of the first turns for both cases start at values close to the averages of the first scenario and the difference between the two cases is rather small. The score gains of the two cases start to diverge strongly only after a few turns, mostly due to a steep increase in the gains for the PROD-case. The growth of the difference between the two cases decreases after turn 7. The mean of the difference in score gains between the two cases for the last 4 turns is equivalent to 30% of a SD.

The two graphs below show the distribution of score gains separated for those students who actively chose schools and for those who stayed at the assigned school. Remember that these values result from the comparison to a hypothetical case where no student was allowed to choose schools in any simulated turn.
Students who actively choose schools fare better than those who stay at the assigned school. Compared to the hypothetical case without choice, choosers loose on average 0.239 points in
the OUT-case, while students who stay loose 0.564 points of average outcome scores. In the PROD-case, choosers gain 2.094 points while students who stay at the assigned school gain 1.575 points. The differences between the two cases are 2.333 points or 20.5% of a SD of outcomes for choosers and 2.139 points or 18.8% of a SD for non-choosers, up from 6.7% and 2.3% in the first scenario. Also note that in the PROD-case, both choosers and non-choosers gain scores. The average gain for non-choosers is equivalent to 13.8% of a SD of outcomes. In the PROD-case of the second scenario, choice therefore really is a “tide that lifts all boats” (Hoxby 2002a).

In the OUT-case, both choosers and non-choosers loose scores when compared to the hypothetical scenario where all students have to attend the assigned school. The average loss of choosers is equivalent to 2% of a SD of outcomes. Although these students care enough about education to attend another than the assigned school, and get access to schools with higher average test results and to peers with a higher ability than in their assigned schools, they slightly loose regarding test outcomes because the schools they choose have on average a lower educational productivity than in a scenario without choice.

Note that this value of 2% is not the loss in scores that the average chooser gets for attending another school than the assigned one. It is the loss that results from comparing the average score at the attended school in the simulation to the score at the assigned school in a hypothetical setting without school choice. The average score gains of choosing versus not choosing for the individual chooser can be found in the NCTT results. These results are equivalent to the effects of winning an entrance-lottery at a popular school as compared to loosing the lottery and having to stay at the assigned school, a setting used in the literature to identify individual academic gains from active school choice. In the PROD-case, the average chooser gains 1.903 points or 16.7% of a SD of outcomes as a result of her decision to attend another than the assigned school. The equivalent gain in the OUT-case is 1.005 points or 8.8% of a SD. These gains are only about half as high as in the PROD-case, but they are still non-negligible and higher than the findings in the literature on the effect of getting access to a popular school. It is however likely that these academic gains would be reduced when the academic costs of attending a more distant school and compensating reactions to loosing a lottery, like additional tutoring or learning support by the parents, are taken into account. The reduction of educational productivity as a reaction to an increase in the academic quality of
the student intake that I presented in 8.8.1 would also likely reduce these gains in the OUT-case in reality\textsuperscript{55}.

The graphs below show the change in average innate ability that results from choice in the two cases, separated for school-changers and non-changers. Remember that these numbers result from a comparison of simulated outcomes to the hypothetical scenario where all students have to attend the assigned school.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{distribution_of_gains_in_ mate Ability_for_choosers.png}
\caption{S2: Distribution of gains in innate ability for choosers}
\end{figure}

In the PROD-case, choosers gain on average 2.517 points or 51\% of a SD of average innate ability of peers as compared to the hypothetical no-choice scenario. In the OUT-case, the gain for choosers is 1.972 points or 40\% of a SD.

\textsuperscript{55} I could not include this effect in the productivity determination function in the simulation, because it is unlikely to exist in the PROD-case and I need to ensure comparability between the two cases.
Non-choosers suffer on average a loss of average innate ability of -1.628 points or 33% of a SD in the OUT-case and -0.843 points or 17% of a SD in the PROD-case. The NCTT gains, the difference in innate peer ability between active choice and attending the assigned school for the individual student when the choice behaviour of all other students is held fixed, is much higher. In the PROD-case, the NCTT gain in innate is 7.531 points or 1.52 SD of average innate ability. In the OUT-case the gain is equivalent to 1.48 SD. That the NCTT gains are higher than the gains from the comparison to the no-choice scenario is not surprising. Students cannot observe average innate ability directly, but other school characteristics like the dropout rate and the share of low-income students are observable and correlated to average innate ability. Students are more likely to leave schools with a low average innate ability, and when exerting choice, they go for schools with a high innate ability. More importantly, students leave schools with a low academic quality measure and choose schools with a high academic quality measure. And the higher the innate ability of a student is, the more likely she is to actively choose schools. This results in high-innate ability students leaving low-quality schools and going to high-quality schools, increasing the differences in innate ability between schools. The NCTT gains measure the effect of the choice between the assigned school and the one the student was predicted to attend, given the choices of the other students and the resulting spread of values for average innate ability.
Summing up the findings for the OUT-case for choosers: For students who actively choose schools based on the measures that are observable in the OUT-case, choice results in a much higher average innate peer ability and a slightly reduced educational productivity. The net effect is an increase in test results by 8.8% of a SD of outcomes. But in the OUT-case, the average educational productivity of schools decreases over time, so that even the scores of choosers are in an average over all turns of the simulation 4.5% of a SD lower than in the hypothetical scenario where choice is not allowed and productivity held fixed.

9.3.2.(3) Share and characteristics of school-changers

The next two graphs show the distribution of the share of school-changers and the average innate ability of a school-changer.

Figure 9-27: S2: Distribution of share of school-changers

In the OUT-case, the share of students who do not attend the assigned school is 47.0% and thus very close to the 47.87% observed in reality in the CPS. The share in the PROD-case is much lower at 26.1%. The distributions of outcomes for the two cases are clearly separate. This difference in chooser-shares between the two cases is even more pronounced than in the first scenario, where the share of changers in the PROD-case is only about 8% below the share in the OUT-case.
As in the first scenario, the drop in the share of students who do not attend the assigned school is mostly due to decreased shares for the student groups who are academically disadvantaged. For African American students, the share is 51% in the OUT-case and 29% in the PROD-case. For Hispanic students, the share of changers is at 19% only half as high in the PROD-case as at 41% in the OUT-case. In the PROD-case, only 9% of low-ability students and 25% of medium-ability students actively choose schools, while the numbers in the OUT-case are 28% and 49%, respectively. As in the first scenario, one part of the explanation for these differences is, that many schools in socially disadvantaged areas are revealed to have a better educational quality than what could be deducted from the meetexceed-measure, and that many seemingly better schools (with high average scores) do not show a high educational productivity. As a result, many students from disadvantaged groups are not willing to accept the increase in commuting time that is necessary to attend a school in a “better” area that is often quite distant.

Contrary to the first scenario, schools can now adapt their level of educational productivity to the competitive pressure that they face. And in the PROD-case, the resulting level of educational productivity is visible and can attract students. Due to the increased incentive effect, many schools adapt their level of educational productivity. Schools that are close to each other face a similar competitive pressure, as they compete with the same other schools and the number, density and quality measures of these competitors are very similar. The levels of educational productivity of two schools that are close to each other are therefore likely to be similar. Students therefore have the choice between their assigned schools and competing schools that are more distant and only show slightly different levels of academic quality, making it less attractive to leave the assigned school.

If the educational costs of changing schools, like time lost on a longer commute, were taken into account, these would therefore be clearly lower on average in the PROD-case, further increasing the advantage in average scores of the PROD-case.
Figure 9-23 shows, that in the PROD-case, the average innate ability of school-changers is higher than in the OUT-case. This is probably connected to the lower share of school-changers in the PROD-case. As innate ability is one of the most important determinants of active school choice, it is likely that those students who still actively choose a school show a higher value in this variable. That school-choosers show a higher average ability in the PROD-case might raise concerns about sorting by ability, for which results are presented in the next subsection.

9.3.2.(4) Sorting by ability

The two graphs below shows the extent of sorting of the students into schools by innate ability. The measure I use is, as for the first scenario, the share of students of a group that would have to change schools, so that all schools would show the same share of this type of student.

Sorting for high ability students is almost identical for both cases at 0.258, meaning that 25.8% of high-ability students would have to be transferred to different schools in order to achieve the same share of high-ability students at all schools. This level of 0.258 is higher than the sorting by ability without choice, based on the assigned schools, at 0.166.
Sorting for low-ability students is lower in the PROD-case at 0.221 than in the OUT-case at 0.255 and in both cases higher than without choice, at 0.134. Thus, school choice increases
sorting by ability, with almost identical effects in both cases for high-ability students and a lower increase in sorting due to choice in the PROD-case than in the OUT-case.

9.3.2.(5) Sorting by race and income
The graph below shows sorting by race for African American and Hispanic students.

Concerning sorting by race for African American students, the average for the PROD-case at 0.728 is hardly above the value without choice, at 0.723. In the OUT-case, sorting is clearly higher at 0.782. For sorting by race for Hispanics, the average for the PROD-case at 0.681 is higher than the value for the case without choice at 0.654 but lower than the average for the OUT-case at 0.705.

The graph below shows the distribution of sorting by race for White students and sorting by income, based on whether students received lunch for free or at reduced prices. I did not include sorting by race for Asian and Pacific Islander students, as they are a rather small minority in the CPS.
The average for sorting by race for White students in the PROD-case at 0.589 is lower than the average for the OUT-case at 0.601. The averages for both cases are higher than the value for the case without choice at 0.545. Sorting by lunch status hardly differs between the three cases, at 0.273 for the PROD-case, 0.266 for the OUT-case and 0.268 for the case without choice.

9.3.2. (6) Capacity utilization of schools

The following graphs show the average capacity utilization of schools over the 10 years of each simulation run. As in the first scenario, I have to consider that changes in the academic quality only become visible when the corresponding cohort has taken the PSAE test in grade 11. As the quality measure in both the PROD-case and the OUT-case are not fixed, the intended quality of the chosen school, based on the observable values when the school choice decision is taken will differ from the actual quality that applies to the cohort of the student. The first set of graphs shows the share of free capacity at the schools that had an educational productivity at or above the average. The left graph shows the intended quality of the school, based on the observable values, the right graph shows the actual values that applied to the cohort of the student.
Based on the quality measure that was observable when the school choice decision was taken, in the OUT-case, the average free capacity at schools with above-average educational productivity shows a flat trajectory with a mean averaged over all turns around 35%. In the PROD-case, the average free capacity at schools with an above-average educational productivity starts around 10% and increases steeply to values around 30% towards the end of the simulated time-span of 10 turns. The averages in the first turns are close to the values for the first scenario, where educational productivity is fixed at the values observed in reality in the underlying base-years.

The reason for this strong increase of the share of free capacity at above-average schools is likely connected to the fact that the share of choosers is much lower in the PROD-case. Schools adapt their level of educational productivity to external incentives in this scenario. Schools that are close to each other face the same competitors and thus similar external pressure and will therefore converge to similar levels of educational productivity. The difference in educational productivity between the assigned school and other close-by schools will therefore be reduced during the simulation to rather small values for most students. As a consequence a setting becomes likely, that might explain the high share of free capacity at above-average schools towards the end of the simulated time-span. In this setting, the level of educational productivity is either for most schools in the area below average or for most schools in the area above average. To attend a school with above-average educational
productivity would then require for a student who lives in an area where schools have below-average productivity a strong increase in distance that will make the change unattractive in most cases. As a result, students remain in their areas, mostly even in their assigned schools and choice does not much change the levels of free capacity that result from the assignment of students to home schools.

Based on the quality measures that applied to the cohorts of choosers, the values for both the PROD-case and the OUT-case vary wildly from turn to turn, especially at the beginning of the simulation. This is to be expected, as in this second scenario, schools can adapt their level of educational productivity to external influences. As seen in 9.3.2.(1), this leads on average to a sizable change in educational productivity, especially in the PROD-case. Parents have to base their choice decision on observations that are visible at the time of the deadline for choice decisions. In the several years that it takes for the actual attended quality to become visible, the quality measures for most schools will have changed strongly. It is likely that, as a result, the quality measure of many schools crosses the mean value of productivity.

The second set of graphs shows the share of free capacity at the schools that had a meetexceed share at or above average. The left graph is based on quality measures that were observable when school choice decisions were taken, the right graph is based on the quality measures that actually applied to the cohort.

![Free share, expected level vs. experienced level](image)

**Figure 9-34: S2: Average share of free places at high-scoring schools by turn, based on expected and experienced level**
Based on the quality measures that were observable when taking the school choice decision, the share of free capacity at above-average schools based on the meetexceed measure shows in the PROD-case a rather flat trajectory around 35%. The free capacity in the OUT-case varies more strongly, starts at 15% and rises to around 25% in the second half of the simulation. The explanation for this increase is likely similar to the one for the measure of educational productivity. But for the meetexceed measure in the OUT-case the effect is much weaker, as educational productivity does not react as strongly to external pressure and can only affect the observable quality measure indirectly.

Based on the values that actually applied to the cohorts, the free capacity of above-average schools based on the meetexceed-share shows for the OUT-case a rather flat trajectory with an overall mean around 22%. The corresponding share in the PROD-case starts around 22% and increases to values above 30%.

9.4. Summary of simulation findings and comparison of outcomes for the two different measures for academic quality

The findings in the two scenarios do not differ strongly, except in the scope of the changes that result from going from the OUT-case to the PROD-case. These changes are in most cases higher in the second scenario, where not only the sorting effect but also the incentive effect is at work. As the differences between the scenarios are similar (except for scope) and the second scenario is much more realistic, the results that are mentioned in this summary will be taken from the second scenario, unless otherwise stated.

In the OUT-case, where academic quality measures of schools are based on the meetexceed output measure, choosers gain from exerting choice, as reflected in the NCTT-outcomes. Choosers gain access to schools that show a higher innate ability than at their assigned school but slightly lose regarding educational productivity. The net-effect is a moderate gain in outcomes, at 8.8% of a SD of outcomes. An academic gain of this size that results from active school choice is hardly found in previous literature. It is however likely to be reduced in reality by three factors that I could not include in my simulation: compensating reactions of parents, commuting costs and a reduction of the effort of schools as a reaction to a higher innate ability of the student intake.

The NCTT gains are the appropriate measure for the gains of individual students from their isolated choice in one turn, but they do not reflect the overall effects of school choice. A comparison to outcomes in a hypothetical scenario where all students have to attend their
assigned schools yields a measure for the overall effects. This overall effect of school choice is a reduction of the educational productivity of the average attended school by about 40% of a SD of educational productivity. The average chooser fares even worse, attending a school with an educational productivity 46% of a SD below the value in the no-choice scenario. This reduction of educational productivity leads to an average reduction in test scores by 4.5% of a SD averaged over all students. Active choosers fare better because they attend schools with a higher level of innate ability than at their assigned schools. But they still lose 2% of a SD of scores as compared to the no-choice scenario. The exodus of mostly high-ability choosers from low-score schools comes at the cost of non-choosers, who face an average innate ability 33% of a SD below the level they would have faced in the no-choice scenario.

In the PROD-case, where academic quality measures for schools are based on their educational productivity, choosers also gain innate ability as compared to their assigned school. But contrary to the OUT-case, they also gain 0.7 SD of educational productivity, resulting in a total score gain of 16.7% of a SD of outcomes, which is about twice as high as the gains in the OUT-case. Moreover, the limiting factors described for the OUT-case are likely lower in the PROD-case. There would be no point in easing effort in response to a higher innate ability of the student intake if educational productivity is observable. Additionally, the share of choosers is lower so that average commuting costs are lower and less students will be denied access at schools to which they applied, resulting in less compensatory measures.

To evaluate the two alternative measures for the academic quality of schools, overall effects of school choice are more important than the NCTT-gains of individual students described above. These overall effects can be identified by comparing simulation outcomes to the hypothetical scenario where all students have to attend the assigned school. Based on this comparison, the effect of school choice on the average educational productivity of the attended school in the PROD-case is an increase by 1.43 SD of educational productivity, as compared to the loss of 0.4 SD in the OUT-case. The gain for active choosers is slightly higher, at 1.46 SD as compared to the loss of 0.46 SD in the OUT-case. These gains are taken from the second scenario and are much higher than in the first scenario. Most of the difference in gains between the two information-cases is therefore mostly due to the incentive effect, that can only be found in the second scenario, and the cross-effects between the sorting effect and the incentive effect. In the PROD-case, choosers also attend schools with a higher innate peer-ability than in the scenario without choice, but the resulting loss of non-choosers at 17% of a SD of average innate peer ability is only half as high as in the OUT-case (33%). Gains in
productivity and changes in the innate peer ability combine to an average total score gain of 1.6 points or 14.2% of a SD of outcomes, which is 18.7% of a SD higher than in the OUT-case. In the last 4 turns, when there was time for the two effects to work on educational productivity, the advantage of score gains of the PROD-case stabilizes at about 30% of a SD of outcomes.

Choosers fare even better than the average, gaining 2.09 points or 18.3% of a SD when compared to the hypothetical scenario without choice, which is an improvement over the OUT-case of 20.5% of a SD. The increase in average educational productivity also results in sizable gains from choice for non-choosers, despite the decrease in innate peer ability that these students suffer. Compared to the no-choice scenario, non-choosers gain 1.57 points or 13.8% of a SD of outcomes, an improvement over the OUT-case of 18.8% of a SD.

These gains are sizable, especially considering that it hardly takes additional resources to achieve them. If student level data are available in such detail as in the CPS, it is only necessary to use a more refined statistical technique to compute the official academic quality measures for schools, to replace the old output-based measures with the new ones and to explain to teachers, principals and parents the intuition behind these new measures. This explanation is no small feat, but compared to the resources that are necessary, for example to reduce average class size by only one student or to provide teachers with significant monetary incentives to alter their behaviour, the costs are almost negligible.

One concern about school choice is that it might lead to an increase in student sorting. Another concern is, that attending another than the assigned neighbourhood school necessitates a longer daily commute that might affect academic outcomes.

In the OUT-case, the share of school-changers is 47% (47.4% in scenario 1), very close to the share observed in reality for the CPS and to which the model was calibrated at 47.87%. In the PROD-case, the share of changers is only 26.1% (39.4% in scenario 1). This reduction is mostly due to lower shares of changers among the academically disadvantaged students: African Americans, Hispanics, medium-ability and especially low-ability students. If the costs of commuting and of not attending the neighbourhood school could be taken into account, the lower share of changers would further increase the advantage in academic outcomes that were already shown for the PROD-case.

56 There is of course a „political“ side to exposing schools in areas with high ability students as less splendid than they are. And it is likely that teachers, principals and maybe even parents at these schools might oppose the introduction of value-added quality measures, provided that they realize the danger for their vested rights. But if a school authority manages to explain to most parents and teachers the concept behind value-added measures, then most parents should support the introduction, as should at least the teachers at schools with a low-ability intake.
Sorting is increased by school choice in all variables for which I could measure it, when compared to the case where all students have to attend their assigned school. Sorting of high-ability students increases from 16.6% to 25.8% for both the PROD-case and the OUT-case\textsuperscript{57}. Sorting of low-ability students increases from 13.4% to 22.1% in the PROD-case and 25.2% in the OUT-case. Sorting of African American students increases hardly in the PROD-case when compared to the no-choice scenario, by 0.5% from 72.3%, but sizably by 5.4% in the OUT-case. Sorting for Hispanic students increases from 65.4% in the no-choice scenario by 2.7% in the PROD-case and by 5.1% in the OUT-case. Sorting of White students increases by 4.4% in the PROD-case, up from 54.5% in the no-choice scenario and by 5.6% in the OUT-case. Sorting by lunch-status hardly differs between the two cases and the no-choice scenario.

Summing up, the PROD-case, as opposed to the OUT-case, leads to higher individual gains for choosers and to a sizable overall increase in educational productivity and scores that also benefits non-choosers. A change to quality measures based on educational productivity instead of the meet-exceed outcome measure also reduces the share of choosers, and thus commuting costs, and reduces sorting based on all analyzed student characteristics except lunch-status and high innate ability, where the differences between the cases are negligible. The overall effects of school choice in the PROD-case, as compared to the no-choice scenario, are a sizable increase in the productivity of the attended school and in student scores for both choosers and non-choosers. These effects take several turns to develop and would therefore be only partially captured if measured shortly after the introduction of choice. The share of students who choose schools and thus create the competitive pressure that induces schools to increase educational productivity is only 26%. Sorting by innate ability increases sizably as a result of school choice, sorting by race is hardly affected for African American students, increases moderately for Hispanic students and sizably for White students. Sorting by lunch-status is not affected.

\textsuperscript{57} Remember that the measure for sorting is the percentage of students of the analyzed group who would have to change schools so that the resulting share at all schools is the same.
10. Conclusions and outlook for further research

The previous chapters have covered a wide array of topics, some of which have been covered more than 200 pages before this one. I therefore briefly summarize the content and insights from the previous chapters below before drawing additional conclusions.

10.1. Summary and conclusions

The proponents of school choice expect several favourable effects to result from choice. One such effect is an overall increase in educational productivity. The mechanism that drives this increase was first described by Friedman (1955).

There are however several implicit assumptions that have to be met for an increase in productivity to result from school choice. The analysis of the existing literature in Chapter 3 showed that in most real world school choice systems not all these assumptions are met. The implicit assumption that parents prefer schools with a high educational quality is confirmed. The literature also shows that teachers in an average school system can increase the educational productivity and did so in several cases when given the right incentives. But I also found that the additional effort that is necessary to increase educational productivity above the level that occurs without competitive pressure is shunned.

Moreover, institutions and regulations in school choice systems play a critical role, and are frequently set in a way so that not all the assumptions that are implicit in Friedmans’ mechanism are met. Even seemingly small misconstructions in the regulations of a school choice system can easily prevent competitive pressure to arise from choice. And even if there is a significant level of competitive pressure, unfavourable settings in the regulations can easily create incentives that induce schools and teachers not to increase educational productivity as a reaction to this pressure, but to cheat instead. In the worst cases of wrongly constructed school choice systems, teachers and principals at a significant share of schools can gain almost nothing from any achievable increase in effort.

Information about school characteristics plays a pivotal role in shaping the incentives and thus determining the reaction of teachers and schools to school choice. Even if all the other regulatory settings are geared for a high impact of school choice on incentives, the wrong kind of information about the academic quality of schools can strongly distort incentives. Parental school choice relies heavily on official quality measures, if these are easily accessible. In most real world school choice systems, these measures are based on raw outcomes. Such output-measures are mostly driven by the average characteristics of the
student intake. Therefore, if observable school quality is based on output-measures, parents often cannot identify the most productive schools. Then, changes in effort, and thus the educational productivity of schools can only have a limited impact on how the academic quality of the school is perceived by parents.

Moreover, output-based quality measures can be rather easily manipulated. If parental school choice and/or rewards and punishments by school authorities are affected by output-measures, incentives for cheating are high. And, as the correlation between output-measures and the educational productivity is low, parents have difficulties to identify the schools that would provide the highest academical benefit to their children. Moreover, if outcome-based quality measures are used, schools with a high-ability student intake can afford a lower effort without appearing to have a low academic quality.

Based on the data from the CPS, I found evidence for several of the potential problems of outcome-based quality measures that I have presented above. Parental preferences in the CPS favour schools with a high academic quality. But as parents cannot observe educational productivity, their choice is based on outcome measures that are mostly driven by average student characteristics. Consequently, the result of active choice is on average a slightly lower productivity than at the assigned school. I also found preliminary evidence that schools deliberately increase the drop-out rate in order to keep weak students out of the centralised tests and that schools reduce their effort, and thus educational productivity, if the innate ability of the student intake increases.

This illustrates that even in a choice system like the CPS, that is otherwise favourable to beneficial effects of school choice, the mechanism that increases educational productivity as a result of school choice can be crippled by one wrong setting, like the use of outcome-based quality measures. This observation is a further indicator on the importance of finding a way to analyze the impact of institutional settings in school choice systems.

As presented in chapter 4, it is however not easy to isolate the impact of individual regulations or external settings in complex systems such as school choice systems. Complications arise mostly from cross-effects between individual regulations and conditions, from the interdependency of decisions among the actors and because a single unfavourable setting can cripple the core mechanisms that drive results. Conventional approaches are not capable of analyzing the impact of individual settings in such complex systems. Standard theoretical model approaches fail, mostly because they become quickly unsolvable if only the
most important cross-effects and interdependencies are integrated. Empirical approaches do not work, mostly because the opportunities for viable identification strategies are very rare. A meta-study approach fails, because there are not that many studies and the information that these contain on regulatory settings and external conditions is in most cases not sufficiently detailed.

A micro-simulation approach has the potential to deal with the complications described above and was used by innovative researchers, such as Epple and Romano (1998), Nechyba (1999) and Fuchs 2007), whose approaches I described. These first simulations used models that were hardly based on real world school choice systems and real world data. Moreover, these models were partial models that were designed to analyze only a limited part of the effects of school choice (chapter 5). The introduction of micro-simulations to school choice was a valuable contribution and the models were detailed enough to analyze those aspects they were meant to analyze. But they were neither detailed enough nor grounded sufficiently in reality to examine the effect of alternative institutional settings on the educational productivity of schools or academic outcomes or the overall effect of school choice. Especially the modelling of the educational production function and of choice behaviour in these models was not detailed enough for such insights.

I developed an own model and programmed a simulation based on this model. The programming was loosely based on a simulation programmed by Fuchs (2007) that I strongly altered and greatly expanded. My model and simulation are based on the school choice system of the CPS and strive to replicate the mechanisms and settings of this real world school choice system (approach described in chapter 6). The CPS provided upon request a dataset that contains information on several school years for several cohorts of students. I completed the dataset by using freely available school-level data on the CPS, computing distances between schools and aggregating additional school-level data from student-level data (chapter 7). Then I estimated all the important functions that drive the decisions of the actors and determine outcomes. Where direct estimations were not possible, I calibrated the model. This calibration is based on observations of real world data of the CPS. Then I populated the model with the same real students and schools that I had used to estimate the model (chapter 8).

The result is a micro-simulation model that replicates the school choice system of the CPS with all relevant actors and their decision functions, institutional settings, external conditions
and mechanisms. Starting from such a model, it is possible to alter individual institutional settings or external conditions and analyze the effect of the alternatives.

This is the great strength of the micro-simulation approach: all the conditions and settings can be controlled and altered at will. Moreover, the simulation can be run for several years, so that potential dynamic effects have the necessary time to unfold. And, unlike equilibrium-based approaches, it is possible to analyze the development of effects over time, not only in a hypothetical equilibrium that might only be reached after decades. Moreover, it is possible to get at least a first impression of the effects of changes that have not been done in reality and could therefore not be observed. However, for the observed effects that result from variations in the settings of a micro-simulation model to have any credibility, it is necessary to be very careful in designing, estimating and calibrating the original model. And it is necessary to consider whether changes in the condition or setting that is analyzed might affect incentives of the actors. If that is the case, decision functions need to be adapted. This kind of adaptation is tricky and has to be done and explained carefully. But if it is done right, the micro-simulation approach allows insights that traditional approaches could not provide so far.

I used such a micro-simulation approach in a first application on a first school choice system in order to examine the effects of value-added quality measures for the academic quality of schools. These value-added measures control for differences in the student intake and thus reveal the actual educational productivity of schools, unlike the output-based measure that was used in the CPS in reality. In a first scenario I assumed passive schools, meaning that schools do not adapt productivity to external pressure. Educational productivity was therefore held fixed at the values observed in the CPS data. This first scenario aims to identify pure sorting effects that occur, if the real educational productivity is made observable.

In a second scenario, I made educational productivity reactive to external pressure. In the preparatory empirical estimations, I had found that educational productivity reacts to free capacity at the own school, probationary status, and the competitive pressure of nearby schools that is a function of the distance, number and quality of these schools. The function for competitive pressure was based on school choice patterns that I observed for students in the CPS. As it becomes much more rewarding for schools to increase educational productivity if this quality measure is made observable to parents, it is necessary to adapt the reaction function of schools that determines educational productivity as a reaction to competitive
pressure. I based this adaptation on the findings of several empirical studies. As there are no conventions or even a precedent for such an approach that I am aware of, I heavily leaned towards a limited adaptation of the decision function of schools and thus towards low effects of the observability of educational productivity. I also limited the actual yearly changes in educational productivity of individual schools to changes that I had observed in the real CPS.

I found that a variation in one institutional setting, the change in the type of the official measure for the academic quality of schools from a raw output measure to a measure based on educational productivity, has sizable effects. In the case where the academic quality measure is based on educational productivity, students are rewarded for active choice by a higher gain in scores, the share of choosers and sorting of students decrease and, compared to the Out-case and to a scenario without choice, schools increase their effort. As a result, average educational productivity and average scores increase by a sizable amount, despite the fact that I leaned heavily towards low effects in the calibration of the model. Moreover, these gains apply to active choosers and non-choosers, to students of all levels of innate ability and of all races, so that school choice in this case really would be a “…tide that lifts all boats..” (Hoxby 2002a, title).

The difference between scenario one and two shows, that only a small share of the beneficial effects of observable educational productivity is due to the pure sorting effect that occurs when students sort themselves into the more productive schools while the educational productivity of these schools is held fixed. Most of the effects of making educational productivity observable come from the changed incentives for schools and the cross-effects between sorting and this incentive effect.

It is important to note, that most of the effects do not appear immediately and it takes several cohorts before most of the effects can be observed. Productivity only starts to increase strongly in the third simulated turn and this increase only becomes observable another 3 years later, when the results of the PSAE-test for the third simulated cohort can be observed. Thus, sizable increases in educational productivity and academic outcomes are only observable 7 to 9 years after the change in the official quality measure.

The main insight gained from this first application of the micro-simulation approach to analyze changes in institutional settings in school choice is as follows:
A change in the measure for academic quality in an existing school choice system from one based on outputs to one based on actual educational productivity strongly increases the beneficial effects of school choice. Such a measure makes the actual contribution of individual schools to the academic success of their students visible. This visibility then makes the mechanism work, that is expected to increase educational productivity as a result of school choice and that was broken in the CPS in the OUT-case. However, a little patience is necessary before a significant share of the full effects can be observed.

Additionally, the simulation of the OUT-case confirmed a suspicion of previous researchers. This suspicion is, that with outcome-based quality measures, parents cannot identify the more productive schools and that, as a consequence, they send their children to seemingly better schools but experience a loss in educational quality. This observation helps explain the low or non-existent academic gains for individual students from active school choice in previous empirical studies.

10.2. Outlook for further research

Chapters 6 through 9 presented a first application of a micro-simulation model that included all main features of school choice and was thoroughly modelled, estimated and calibrated based on real world data, in order to identify the effects of a change in an institutional setting. This approach could be expanded, both for the analysis of changes in other institutional settings or for changes in exogenous conditions for the CPS and, in the long run, for a deeper understanding of school choice in general.

With the model already estimated and calibrated, the impact of changes that do not affect the incentives or information of the actors strongly could be analyzed with a moderate additional effort. One example would be the effect of a higher density of schools. Such a density could result, for example, from splitting the largest schools into two schools with a separate curriculum, staff and principal. Alternatively, the distances between schools could be multiplied by a factor $<1$, as an analysis of the changes of the effect of school choice depending on the density of the population. The higher density of schools would increase competition, as there would be more schools within commuting distance. The incentives and thus the decision functions of the actors should not change fundamentally if the distance of
competing schools changes. This change could be reflected in the model by only changing the variable values that go into the decision functions.

Unlike in the example above, the analysis of alternative settings for most exogenous conditions and institutional settings is likely to result in a change of incentives for at least one group of actors and thus requires an adaptation of decision functions. An example for such an institutional setting that could be incorporated rather easily would be the impact of an increased support for transportation. This support could be an extension of school-bus routes or higher subsidies for fares. Such changes could be incorporated into a model by a decrease in the discounting factor for distances in the parental school choice decision function and small adjustments to the distance-weight in the computation of competitive pressure for schools.

Changes in other institutional settings require a more extensive analysis of the impact on the incentives of actors. An example are all limitations to the choice options of parents. If, for example, the free capacity in the school system was reduced, then fewer students could leave unpopular schools. This would reduce the threat of student loss for unpopular schools and should be reflected by a less responsive productivity determination function. If the share of students who are allowed to leave an individual school was limited, like in several real world school choice systems, this would further reduce the threat to low-quality schools, as their student losses would then have an upper bound. Moreover, increases in quality that do not reduce the share of would-be choosers to a number below the limit would have no effect on actual student numbers.

Among the changes in institutional settings that would distort incentives most heavily are changes that selectively limit the right to choose or the admission to schools based on student characteristics. If school choice is limited to students from failing schools, for example, then schools would expect the average chooser to have an innate ability that is below average. In this case, schools would be less eager to attract choosers than if choice is free to all students, as then choosers usually have an innate ability above average. This relative unattractiveness of choosers would be increased further, if the measure for academic school quality was based on outcomes. Then, attracting many school-changers from failing schools might reduce the quality measure of a school. In this case, being attractive to school-changers might even be undesirable for schools. Selective student admission would have an opposite effect. If admission to schools was not determined by a lottery as in the CPS, but schools were allowed to pick their freshmen, it would be beneficial for schools to be attractive to students. The schools could then cream-skim and reject those applicants who could hurt the score average.
Changes in institutional settings that selectively limit choice or admission would necessitate a very careful adaptation of the decision functions of the actors. Such adaptations would also need to be based on extensive empirical foundations in order to yield reliable results.

In order to gain insights into some potential changes, it would even be necessary to extend the model. In order to analyze, for example, the effect of an introduction of additional school-bus lines that transfer students from disadvantaged areas to high-quality schools, it would be necessary to refine the way distances are modelled. The current use of distances between schools as the crow flies would be no longer appropriate, as some schools could then be reached more easily than implied by the simple distance. Ideally, the distances would be replaced with a measure based on travelling times.

Based on the approach developed in this dissertation, it is possible to analyze the effects of a wide range of changes to regulatory settings and of changes in exogenous conditions, provided that there is sufficiently detailed data to estimate the crucial functions that drive the system and provided that only one setting or external condition is changed at a time. Moreover, the simulation of several turns would make it possible to incorporate dynamic effects and to get insights into the development of effects over time. The results should be quite reliable, even if a potentially necessary adaptation is done, provided that this adaptation can be based on sufficient empirical evidence from other school choice systems and is done carefully.

Once a regulatory setting or external condition has been analyzed with the approach described above for several school choice systems, it would become possible to get an insight into the impact of this setting or condition in general, and how this impact depends on other settings and conditions. These insights could help explain the diverse findings in the literature about the effects of school choice.

In the long run, once potential changes to the most important regulatory settings and external conditions have been analyzed and the decision functions of the actors have been estimated for several school choice systems, the thus improved understanding should help to reconcile the often diverse and contradictory empirical findings on school choice. This understanding should also help to identify a range for expected effects of school choice on different types of students.

It should also become possible to predict, at least roughly, the effects of introducing school choice in a school-system where there is currently no choice. It would even be possible to analyze the effects that several potential shapes of the new school choice system would have
over a prolonged period of time, a valuable tool for school authorities when deciding whether to introduce school choice and how to design a potential school choice system.
Appendix 1: On the software STATA and replication of results:

The simulations and preparatory estimations in the second part of this dissertation were done using the software STATA (Stata 10), run under Windows XP, SP 2. I have included all the datasets and command files that are necessary to replicate these simulations, the preparatory estimations and data analysis on the accompanying DVD. These files also contain additional insights on several topics for the interested reader. For the replications to work and the additional insights to be available, the folder “Daten” must be copied to the location “C:\Daten”. Any student-level dataset must only be used for the purpose of grading this dissertation. For any other use it is necessary to consult the conditions under which these data were supplied by the CPS. It is likely, that any other use has to be authorized by the Chicago Public Schools. This dataset is therefore only included in the DVDs for those copies that are handed in for the purpose of grading the dissertation.

To replicate the simulations, it is also necessary to install several commands that are not included in a fresh install of STATA 10. These commands are included in the packages “uvis”, “mstore” and “duncan” and easily installed via the “findit” command. As both my version of STATA and the additional commands are rather seasoned by now, I have included both on the DVD. For the version of STATA that I used, copy the folders “Stata10” and “ado” on the DVD directly onto “C:\”, to the locations “C:\Stata10” and “C:\ado”. This transfers the STATA version that I have used as well as all additional commands.

For the case of passive schools (scenario 1 in chapter 9.2.1) the data preparation is done by running: “C:\Daten\Stata CPS Data\Data Preparation\Prepare Dataset Differentiate.do”

For the case of schools that adapt their educational productivity to external pressure, the data preparation is done by running:

“C:\Daten\Stata CPS Data\Data Preparation Schoolmax\Prepare Dataset Schoolmax 2.do”

The simulations are replicated by running, for the case of passive schools, the do-file:

“C:\Daten\Stata CPS Data\A Master Shell Updating Differentiate.do”

For the case with adapted educational productivity, run:

“C:\Daten\Stata CPS Data\Simulation including schoolmax\A Master Shell Schoolmax.do”

These files are preset to run simulations of 10 turns for both information-cases 50 times. On an average personal computer, the running time should be several days of continuous computation. The file that holds the outcomes for all simulation runs of the corresponding command file are stored in "C:\Daten\Daten CPS\MasterUpdatingTest.dta” for the case of
passive schools and "C:\Daten\Daten CPS\MasterSchoolmax.dta" for the case with adapted educational productivity. The outcomes of 200 simulation runs for each scenario, which I used to get insights and generate the outcome graphs of the dissertation, are stored in "C:\Daten\Daten CPS\Outcomes" and are called “MasterSchoolmaxActiveGesamt200” for the scenario 2 and “MasterSchoolmaxPassiveNCTT200” for scenario 1.

The folder “C:\Daten\Stata CPS Data\Graphs and Tables” contains do-files that generate most of the graphs and tables used in this dissertation and provide additional insights for the interested reader. I refer to several of these files in the text of the dissertation for additional insights on topics that are covered in the corresponding chapters.

I have commented the do-files extensively, so that other users should be capable of understanding of what I intend to do in these files and how I use STATA to achieve these ends.
Appendix 2: School level data

The school level data on enrolment, racial composition, truancy rate etc. are provided by the CPS in one file that contains these data for the years 1989 through 2006 and for all schools in the CPS. I harvest the data for the schools in my sample from this file and store them in matrices using the command “mstore” to have them available when using the dataset that contains the student level data. I therefore have no need for a file that contains the school level data for only the schools that are included in my simulation or the preparatory estimations.

The original data file can be accessed on the accompanying DVD as an Excel-file at:
“Daten\Stata CPS Data\Data Preparation\Schoolinfo 89-06 usable”

Or as a STATA data file at:
“Daten\Stata CPS Data\Data Preparation\School Data 89-06 usable”

The schools that were used in the simulation have the following values for “unit-number”:

<table>
<thead>
<tr>
<th>1010</th>
<th>1030</th>
<th>1070</th>
<th>1110</th>
<th>1140</th>
<th>1150</th>
<th>1160</th>
<th>1200</th>
<th>1210</th>
</tr>
</thead>
<tbody>
<tr>
<td>1220</td>
<td>1230</td>
<td>1240</td>
<td>1250</td>
<td>1270</td>
<td>1280</td>
<td>1300</td>
<td>1310</td>
<td>1320</td>
</tr>
<tr>
<td>1330</td>
<td>1340</td>
<td>1350</td>
<td>1360</td>
<td>1380</td>
<td>1390</td>
<td>1400</td>
<td>1410</td>
<td>1420</td>
</tr>
<tr>
<td>1430</td>
<td>1460</td>
<td>1470</td>
<td>1480</td>
<td>1490</td>
<td>1510</td>
<td>1520</td>
<td>1530</td>
<td>1540</td>
</tr>
<tr>
<td>1550</td>
<td>1560</td>
<td>1570</td>
<td>1580</td>
<td>1590</td>
<td>1600</td>
<td>1610</td>
<td>1620</td>
<td>1630</td>
</tr>
<tr>
<td>1640</td>
<td>1670</td>
<td>1680</td>
<td>1710</td>
<td>1790</td>
<td>1820</td>
<td>1830</td>
<td>1840</td>
<td>1860</td>
</tr>
<tr>
<td>1870</td>
<td>1880</td>
<td>1890</td>
<td>1930</td>
<td>6620</td>
<td>6630</td>
<td>7370</td>
<td>7660</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3: Deutsche Zusammenfassung

Die Arbeit untersucht den Effekt von Rahmenbedingungen und institutionellen Regelungen im Bereich der Schulwahl mit Hilfe von Mikrosimulationen, einem Ansatz, der in dieser Dissertation weiter entwickelt und auf eine erste Fragestellung angewendet wird.

freier Schulwahl effektiv wieder zu Nachbarschaftsschulen mit nur wenigen freien Plätzen für Schüler von außerhalb.

Welche Anreize durch Schulwahl entstehen hängt insbesondere stark davon ab, welche Qualitätsmaße Eltern zur Verfügung stehen bezüglich der akademischen Qualität der einzelnen Schulen. Diese Maße haben einen starken Einfluss auf die Schulwahlentscheidung. Meist beruhen Maße für akademische Qualität direkt auf Testergebnissen, welche stark durch die Charakteristika der Schüler bestimmt werden. Mittel- und Oberschichtkinder deren Muttersprache die Landessprache ist, erzielen z.B. im Schnitt bessere Testergebnisse als Migrantenkinder aus einem Arbeiterumfeld. Sind die akademisch relevanten Eigenschaften in der Schülerschaft systematisch ungleich verteilt und werden diese Unterschiede bei der Qualitätsschätzung nicht berücksichtigt, werden Durchschnittsnoten stärker von der Zusammensetzung der Schülerschaft bestimmt als von der tatsächlichen Bildungsproduktivität der Schule. In diesem Fall können Eltern die guten Schulen nicht erkennen und fragen diese nicht vermehrt nach. In der Folge lohnt es sich für Schulen nicht, die Produktivität als Reaktion auf Wettbewerbsdruck zu erhöhen und es ist kein Effekt von Schulwahl auf die durchschnittliche Schulqualität zu erwarten. Auch lassen sich Qualitätsmaße, die direkt auf Testergebnissen beruhen leicht manipulieren, um eine höhere Qualität der Schule vorzutäuschen. In der Literatur finden sich viele Wege, auf denen dies möglich ist und Nachweise für durchgeführte Vortäuschungen.

Es ist möglich, Qualitätsmaße zu generieren, die die Unterschiede in der Schülerschaft berücksichtigen und so die tatsächliche Bildungsproduktivität der Schule sichtbar machen. Hierzu werden für einzelne Schüler Vorhersagen über Testergebnisse berechnet, welche auf vorherige Schätzungen und den Eigenschaften der einzelnen Schüler beruhen. Diese Werte werden mit den tatsächlichen Testergebnissen verglichen und die Abweichung des erzielten Testergebnisses vom vorhergesagten Testergebnis berechnet. Durchschnittswerte dieser Abweichungen über alle Schüler an einer Schule zeigen, wie viel die Schule den Schüler über das zu erwartende Maß hinaus beigebracht hat.

Würden solche Maße verwendet, könnten Eltern anhand der offiziellen Qualitätsmaße die tatsächliche Bildungsproduktivität der Schule erkennen, statt hauptsächlich Informationen über die Schülerzusammensetzung zu erhalten. Dann könnte der von Friedman beschriebene Mechanismus funktionieren und Schulwahl zu Produktivitätsgewinnen führen.

Es ist daher für das Verständnis von Schulwahl wichtig, den potentiellen Effekt einer Einführung solcher Qualitätsmaße zu untersuchen. Allerdings ist es schwierig, mit


In dieser Arbeit entwickle ich ein eigenes Modell und programmier eine darauf basierende Mikrosimulation. Ausgangspunkt ist eine Programmierung von Thomas Fuchs. Diese ändere und erweitere ich stark, so dass vermutlich keine einzige Zeile des Programm-Codes unverändert enthalten ist. Aber die Arbeit von Fuchs zeigt Wege auf, wie manche Elemente einer Mikrosimulation in STATA programmiert werden können, was eine große Hilfe war. Dieses Modell beruht auf Daten aus einem real existierenden Schulwahlsystem, den Chicago Public Schools (CPS). Mein Modell ist wesentlich vollständiger als die bisherigen Mikrosimulationsmodelle. Es enthält umfangreiche Daten über reale und damit heterogene Schüler und Schulen sowie Entscheidungsfunktionen für alle relevanten Akteure und bildet


In dieser Dissertation wende ich diesen weiter entwickelten Mikrosimulationsansatz auf eine erste Variation in einer institutionellen Regelung auf ein erstes Schulwahlsystem an. Auf den Wechsel im offiziellen Maß für die akademischer Qualität von Schulen im CPS von einem direkt auf Testergebnissen beruhenden Maß (OUT-Fall) auf ein Maß, dass die Charakteristika der Schüler berücksichtigt und den tatsächlichen Beitrag der Schule zum Bildungserfolg aufzeigt (VA-Fall). Dieser Wechsel hat starke Auswirkungen. Im VA-Fall werden Schüler im Vergleich zum Out-Fall für aktive Schulwahl mit einem höheren Gewinn an Testergebnissen

Der Vergleich zwischen zwei Szenarien zeigt, dass nur ein kleiner Teil dieser Effekte auf Sorting zurückgeht, also darauf, dass die sonst freien Plätze an den besten Schulen gefüllt werden. Der Großteil geht auf die Steigerung der Bildungsproduktivität in Folge von Wettbewerb zurück sowie auf Kreuzeffekte zwischen Sorting und Wettbewerbsdruck. Wichtig ist auch, dass der Großteil der Effekte nicht sofort sichtbar wird. Die Produktivität steigt erst nach ca. 3 Jahren deutlich an, wenn anhand der Testergebnisse im dritten Jahr der Highschool die Auswirkungen von Produktivitätsänderungen im ersten Jahr der Simulation erstmals sichtbar werden. Und die Auswirkung dieses Anstieges wird erst sichtbar nach 7 bis 9 Jahren, wenn wiederum die Testergebnisse derjenigen Schüler sichtbar werden, die in ihrer gesamten Highschool Zeit von der höheren Produktivität profitiert haben.

Darüber hinaus bestätigt meine Analyse eine Vermutung aus vorherigen Studien: Wenn das Maß für die akademische Qualität von Schulen direkt auf Testergebnissen beruht, können Eltern die tatsächliche Bildungsqualität kaum erkennen und schicken ihre Kinder nur auf scheinbar bessere Schulen. Da diese aber keine höhere Bildungsproduktivität aufweisen, resultiert der Schulwechsel nicht in besseren Testergebnissen. Diese Beobachtung hilft zu erklären, warum die meisten empirischen Studien den erwarteten positiven Effekt von aktiver Schulwahl nicht finden.

Die beschriebene Anwendung ist nur eine erste Anwendung des erweiterten Mikrosimulationsansatzes auf eine Änderung in einem existierenden Schulwahlsystem. Mit weiteren Anwendungen ließen sich noch wesentlich mehr Erkenntnisse erzielen.
Mit den hierbei gewonnenen Erkenntnissen über Regelungen und Rahmenbedingungen sollte es möglich sein, die oft widersprüchlichen empirischen Erkenntnisse zur Schulwahl vereinbar zu machen und die tatsächlichen Effekte in individuellen Schulsystemen zu verstehen. Auch wäre es möglich, Aufsichtsbehörden in Schulsystemen dahingehend zu beraten, Schulwahl so zu gestalten dass sie tatsächlich die gewünschten Ergebnisse erzielt statt kaum Effekte zu zeigen oder die Schulen und Lehrer zu Täuschungsmanövern zu motivieren oder geradezu zu Täuschungen zu zwingen.
Allerdings ist große Vorsicht geboten, wenn relevanten Rahmenbedingungen oder Regelungen variiert werden. Immer dann, wenn die Anreize von Akteuren oder ihr Entscheidungsspielraum verändert werden, muss darüber nachgedacht werden ob die untersuchte Variation die Entscheidungsfunktionen beeinflusst. Ist dies der Fall, müssen diese Funktionen angepasst werden. Hierbei ist große Vorsicht geboten um zu verhindern, dass beobachtete Effekte auf unzureichend begründeten Anpassungsentscheidungen beruhen. Und es ist nötig, den Adressaten der Untersuchung diese Anpassungen ausführlich zu erklären und plausibel zu machen.
List of References


CPS IEP, Chapter 4 of this guide describes different types of students that are eligible for IEP http://www.specialeducationadvisor.com/school-districts/chicago-public-schools/, 20.06.2012

CPS School Data, a homepage that provides downloadable school data, to be found at: http://research.cps.k12.il.us/cps/accountweb/Reports/download.html, 28.01.2011

Information on student characteristics from 1989 to 2006 is in the file schlinfo.zip

An example for a gains statistic is itbs.gains.zip


CPS Value Added, an explanation of the value-added measure used by the CPS http://research.cps.k12.il.us/export/sites/default/accountweb/Research/ValueAdded/FAQ_for_Lesson_1.pdf, 28.01.2011

CPS NCLB, a set of frequently asked questions on the No Child Left Behind policy in the CPS, explaining that students of failing schools are allowed to apply to other schools http://www.cps.edu/Programs/DistrictInitiatives/NCLB/NCLBTransferProgram/Documents/NCLBFAQs.pdf, 20.06.2012

Cullen, J. and Jacob, B. (2007) Is Gaining Access to Selective Elementary Schools Gaining
Ground? Evidence from Randomized Lotteries, *NBER Working Paper 13443*


Choice Behaviour in Public School Choice, *NBER Working Paper 12995*


Hoxby, C. (2002a) School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?), *NBER Working Paper 8873*


Stata 10, StataCorp. 2007. *Stata Statistical Software: Release 10*. College Station, TX: StataCorp LP


WBEZ (2011) Mapping 10 years of school closures