ABSTRACT

XU, ZHIYI. Demand Estimation of Schools Under Manipulable Mechanisms. (Under the direction of Robert Hammond.)

My dissertation is about demand estimation under manipulable school choice mechanisms. There are three popular mechanisms in practice: the *Boston Mechanism* (BM), *Deferred Acceptance Algorithm* (DA), and *Top Trading Cycles Algorithm* (TTC). BM is famous for being manipulable, and DA and TTC are strategyproof. I study the strategic behaviors, efficiency, and fairness under different mechanisms.

In the first chapter, I simulate BM, constrained DA, and unconstrained DA under three different information environments: (1) with complete information about cutoffs and lottery, (2) agents don't know lottery, and (3) agents don't know decimal part of cutoffs. The results show that there's heavy manipulation under BM, and that the truth-telling rate is subject to different information environments. I also study how assignment priority and correlation among preferences affect truth telling rate. Under both BM and constrained DA, the simulation with higher correlation among preferences of schools has more manipulation and lower truth-telling rate, under all the environments. I then compare the assignment probability using the cutoff approach to the approach of counting assignment. An important conclusion is that the cutoff approach works better in practice, especially under manipulable mechanisms or in school districts with many schools and/or complicated priority structure.

In the second chapter, I build the structural model to estimate demand of magnet schools in Wake County Public School System (WCPSS). I analyze data from the academic year 2014-2015, which is the last year that BM was used for magnet school application. I first estimate demand with a homogeneous sophistication model assuming all students are sophisticated. I then expand the model to a heterogeneous sophistication model that allows for both sophisticated and sincere students. I use a unique login data to help me identify the sophistication level of students, and estimate demand given the sophistication level. Explaining variables include both individual level characters whose estimates vary by schools and school level characters whose estimates are constant across schools. The results match anecdotal evidence concerning preferences.

With the revealed preference, I provide information about strategic behaviors in real life. I also conduct several counterfactual analyses on different designs of a school choice mechanisms to examine their effects on efficiency and fairness. The efficiency differences under different mechanisms are small. Meanwhile, TTC is the least fair mechanism, followed by BM and DA. Having a constrained list does not seem to play a big role in efficiency but does increase justified envy. Having reserved seats has mixed effects on efficiency but less justified envy. In contrast, single versus multiple tie-breaking lotteries do not seem to differ much in terms of either efficiency or

fairness. Finally, sophisticated students gain under a manipulable mechanism at the expense of sincere students, on average. Therefore, a strategyproof mechanism would reduce the disadvantage of sincere students. The results suggest that a school choice mechanism must be carefully designed along several dimensions.

In the third chapter, I solve the computation difficulty related to covariance estimation in the Multivariate probit (MVP) model, by introducing Gaussian graphical models (GGM). I adopt the recently developed Bayesian graphical Lasso model to MVP model estimation. I adapt [Wan12]'s Bayesian adaptive graphical Lasso by building a new Gibbs Sampling scheme to estimate MVP models with a penalized covariance matrix. The methodology allows estimation in school systems with up to hundreds of schools, which is an important improvement relative to the proceeding model that has computational difficulties above around 20 schools. I show that this new methodology can estimate the MVP model efficiently without introducing extra computation burden. I further discuss the limitation of the Bayesian graphical lasso and possible solutions using the variational Bayesian inference.

Demand Estimation of Schools Under Manipulable Mechanisms

by
Zhiyi Xu

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Economics

Raleigh, North Carolina

2018

APPROVED BY:

Umut Dur	Thayer Morrill
Subhashis Ghoshal	Robert Hammond Chair of Advisory Committee

DEDICATION

To my loving parents and encouraging husband without whom this work would not have been possible.

BIOGRAPHY

Zhiyi Xu was born October 6, 1988 in Nei Mongol, China. She is also known as Alicia. She attended Baotou No.1 Middle School and Baogang No.1 High School, where she developed her passion of math and science.

She enrolled in Nankai University in 2007 with a major in International Economics and Trade. She graduated in 2011 with a Bachelor of Science in Economics, and then she moved to the United States for graduate study. She attended Vanderbilt University in Nashville, TN for the graduate program in economics, and received her Master of Arts in 2013. In 2013, she was admitted to the doctorate program in Economics at North Carolina State University, and she received her Doctorate degree in May 2018.

During her last two years of Ph.D. study, Zhiyi worked as a year-round intern at SAS Institute Inc. in Cary, NC. She received several offers prior to graduation and accepted the offer from Bates White Economic Consulting. Zhiyi plans to work for Bates White after graduation and devote herself to ensuring the competitiveness of markets.

ACKNOWLEDGEMENTS

First, I would like to express my special appreciation and thanks to my advisor Professor Robert Hammond. He always gives me great freedom to pursue what I want to do, from coursework to dissertation, and from research topic to career choice. I appreciate all his generous contributions of time, ideas, and guidance to make my Ph.D. fun and productive. The research process is not always as smooth as we hope, but his encouragement and persistence helped me overcome many obstacles.

I would also like to thank Professor Subhashis Ghoshal, Professor Thayer Morrill, and Professor Umut Dur for serving as my committee members. I am very thankful for all the great comments and suggestions they provided to improve my research and presentation. I also want to thank many other professors in the Economics Department at North Carolina State University, as they helped me at different stages of my Ph.D.: Professor Xiaoyong Zheng, Professor Mehmet Caner, Professor Melinda Morrill, Professor Walter Thurman, Professor Denis Pelletier, and Professor Giuseppe Fiori.

Lastly, I would like to thank my family for all their love and encouragement. Nothing can express how grateful I am for my parents. This work would not have been possible without their generosity, patience, and support. My husband Tim has also been patient and loving with me since the first day we met, which happens to be a time I felt lost in the Ph.D. program. He has taken over many housework chores to reduce my burden, and he gave up his dream job to support my career. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

CONSISTENT ESTIMATION OF ASSIGNMENT PROBABILITIES

1.1 Introduction

Price is the most common tool for resource allocation in daily life. However, it is not appropriate in many cases, such as marriage, kidney transfer, and public school assignment. This is where matching mechanisms come into play. School choice mechanisms are widely used to assign students to public schools across the world. Students are asked to submit a ranking for schools where they want to attend. The assignment is run based on ranking submitted, priority structures, and the mechanism chose by the school district. There are three popular mechanisms in practice: the *Boston Mechanism* (BM), *Deferred Acceptance Algorithm* (DA), and *Top Trading Cycles Algorithm* (TTC).[1] Under BM, a student who has been rejected by his first k ranked school will be consider for his $(k+1)$th ranked school. However, he will be considered at this school after those students who rank this school higher than $k+1$. So students may lose their priority at a school if they do not rank it highly. This feature causes students to strategize their submitted rankings of schools, that is, BM is a *manipulable* mechanism. In contrast, DA and TTC are *strategyproof* mechanisms, meaning agents have no

[1] BM is used in Barcelona, Beijing, Charlotte-Mecklenberg, Denver, Miami-Dade, Minneapolis, Seattle (pre 1999 and post 2009), and Tampa-St. Petersburg [AS16]; DA is used in New York City, Ghana, various cities in England, Boston, and Seattle (1999-2008) [AS16]; and TTC was used for 2011 in New Orleans but they switched to DA in 2012 [Abd17]. The Wake County Public School System in North Carolina used BM before 2015 and switched to DA in 2015 [Dur17a].

incentive to manipulate rankings and therefore report their true preferences. Manipulation is one of the biggest concerns that pushed many school systems to switch from BM to DA or other non-manipulable mechanisms. Even Boston Public Schools (BPS) where BM was first used to assign K–12 students changed to DA in 2005. I study the Wake County Public School System (WCPSS) in North Carolina, which also switched to DA in 2015. However, the strategyproof is often achieved at the cost of losing efficiency, as [Abd11] find out that BM Pareto dominates DA. To compare the gains and losses between DA and BM, it is important to get an accurate estimation of demand for schools. In addition to a welfare analysis, demand estimation for schools is also important for school districts and policymakers to better understand the popularity of schools and how to allocate resources. Demand estimation is much more difficult under manipulable mechanisms than under DA. I estimate demand using data from WCPSS during the period when it was using BM. This approach provides a framework for additional demand analysis of other years of WCPSS data and of data from other districts.

Due to estimation difficulties, there have not been many empirical works about demand estimation under manipulable mechanisms, especially that very few papers use structural models. Knowing that agents have incentives to manipulate their reports, it is not appropriate to take reported rankings as true preferences. Many laboratory experiments have shown that the rate of truth-telling is much lower under BM than under DA, such as [CS06], [PP08] and [Cal10]. Some recent empirical papers also demonstrate the strategic behavior of agents under BM ([AS16], [He16], [Hwa16], and [Cal16]). Under BM, many agents may avoid putting a competitive school on the list, even if they prefer this school more. If we take reported rankings as truth, we may underestimate the preference and popularity of some more competitive schools and overestimate the preference of some less competitive schools. Failing to consider manipulation will lead to biased estimation, and any welfare analysis or policies that are made based on estimation results will be misleading as well

It's commonly believed that rational agents strategize rankings in order to maximize their expected utilities. It's similar to traditional discrete choice model with random utility. The differences are (1) there's uncertainty in the process of school mechanism and (2) students choose an ordered list of options, instead of choosing only one choice. If we could get the assignment probabilities of each student in each school for each feasible ranking, then the problem is reduced to a discrete choice model where the maximizing subject is a linear combination of different schools with the highest expected utility. Therefore, the first step of demand estimation is to estimate assignment probabilities.

There are two approaches to calculating assignment probabilities in literature. [AS16], [Kap17], and [Hwa16] provide a consistent estimation of probabilities using cutoffs.[2] Both [He16] and [Cal16]

[2]They have different definitions for cutoffs, but of the idea is similar. [Hwa16] defines cutoffs as the minimum lottery needed in a critical round (the round where the assignment of a student would depends on his lottery number) for a student. The cutoffs therefore vary by students. The cutoffs in [AS16] involve assignment mechanism, priorities, and

get their assignment probabilities by counting the numbers of students who get assigned to each school conditional on each ranking and priority type. A big problem with this approach is that some possible rankings may not be observed in the data for certain priority types and therefore we cannot get the probabilities for these rankings. Even if all possible rankings were submitted, they may not be submitted enough to get a consistent estimator. More schools in the application system, more possible rankings there are. With 23 elementary schools in my data, there are 11,156 possible rankings, but there are only 2026 applicants. The approach of [He16] and [Cal16] cannot be applied to my data. [AS16] provide a convergence condition for estimates of assignment probabilities. In particular, they prove that a very wide class of mechanisms all satisfy this convergence condition and they name it *Report-Specif Priority and Cutoff Mechanisms*.

In this chapter, I first explain what the *Report-Specif Priority and Cutoff Mechanisms* is with a counter example. I then demonstrate a more practical way than [AS16] to construct eligibility scores and cutoffs, which are used to estimate the assignment probabilities. Under the cutoff approach, I simulate BM and DA. When considering DA, I consider a situation where students can submit a complete list of schools (unconstrained) or only a subset of schools (constrained). I study these mechanisms in different informational environments and with different priority structures. Assuming all students are fully sophisticated and maximize their expected utility. I compare equilibria, cutoffs, and manipulations under these different scenarios. While there is more manipulation under BM, simulation results show that the basic pattern of truth-telling rates under different scenarios hold for both constrained DA and BM. There is more truth telling with partial information than with complete information about cutoffs and tie-breaking lotteries, because agents will rank a competitive school first as long as they believe their assignment probability is high enough. With full knowledge of cutoffs and tie-breaking lotteries, agents avoid competitive schools, because they know they cannot get in. The simulations with higher correlation among preferences of schools has less truth-telling under all the environments. Higher correlation means popular schools are even more competitive, therefore more students avoid them. In the last section of this chapter, I compare the method of estimating assignment probabilities using the cutoff approach to the approach of counting assignments. It shows that the cutoff approach has many practical advantages for empirical work, especially under manipulable mechanisms or in school districts with many schools.

1.2 Report-Specific Priority and Cutoff Mechanisms

[AS16] propose a convergence condition by defining the mechanism as a function that would assign a unique assignment to an agent based on the rankings, priority type, and tie-breaking lottery of the agent. This convergence condition allows us to consistently estimate of the assignment probabilities

lottery and are the same for all the students. In this paper, I use the definition of [AS16].

when the samples are generated from a sequence of type-symmetric strategies. However, it is not easy to verify a mechanism satisfies the convergence condition. To solve this, [AS16] categorize a class of mechanisms that would satisfy the convergence condition. They name it as *Report-Specif Priority and Cutoff Mechanisms*. They point out that any mechanism in this class could get consistent estimates of assignment probabilities using *market clearing cutoffs*. For any mechanism in this class, an agent can be assigned an eligibility score at each school through the function f, based on the ranking submitted, the priority type, and the tie-breaking lottery. Market clearing cutoffs are similar to market price. When a student is being considered at a school, if his eligibility score is higher than the cutoff of this school, then this student will be assigned there, and rejected other wise. As long as the the capacity of a school is exhausted, the cutoff is strictly positive. If there are still seats left in a school, the cutoff is zero. Similar to market price in a perfectly competitive market, each individual cannot decide school cutoffs. Individuals take cutoffs as given and make decisions accordingly. The key restriction of assigning eligibility scores is that the function f does not depend on the reports and priorities of other agents, meaning a student's assignment does not depend on others' ranking or assignment. Violating this restriction is the most common reason for a mechanism fails to be a Report-Specific Priority and Cutoff Mechanism, such as the Cambridge Mechanism in [AS16] and TTC. To better illustrate the concept of Report-Specif Priority and Cutoff Mechanisms, I use the Cambridge Mechanism as a counter example and explain why it is not in the class.

Cambridge Mechanism is a modified version of BM. The assignment algorithm is same as BM, except that there are two types of capacities in Cambridge, the capacity of a program and the capacity of a school. Each school in Cambridge has two programs, a federal lunch program (FLP) and a paid lunch program (PLP). Each school sets capacities for both the lunch programs and the school. A student could only be assigned to a program in a school if there are still seats for both the school and the program he applies for. The Cambridge is same as BM if the sum of capacities of the two programs is smaller or equal to the capacity of the school, and hence it is a Report-Specif Priority and Cutoff Mechanism. However, the sum of capacities in the two program is bigger than the capacity of the school in Cambridge. Furthermore, the Cambridge algorithm is not a Report-Specif Priority and Cutoff Mechanism in this case, as shown in the example below.

Example *Cambridge Mechanism is not a Report-Specif Priority and Cutoff Mechanism*

There are 4 students, s1, s2, s3, and s4. The first two students apply to FLP and the later two apply to PLP. The capacity of school A is 3, but the capacity of FLP in school A is 2 and the capacity of PLP in school A is 2. The lottery ranking, from highest to lowest, of the four students is : s4, s3, s2, and s1.

- Scenario 1: s1, s2, s3 ranking A first, and s4 rank A second.
    - s1, s2, s3 get into A. s4 get rejected by A.

$Cutoff_{PLP,A} = 0$ because PLP has one seat left. The eligibility score of s4 in school A $e_{s4,A} \leqslant 0$ because s4 get rejected by A and $Cutoff_{PLP,A} = 0$.

- Scenario 2: s1, s3 ranking A first, s2 and s4 rank A second.
  - s1, s3 get into A. s2 and s4 get considered by A in second round and s4 get in because he has higher lottery number.

    $Cutoff_{PLP,A} = e_{s4,A} > 0$ because PLP has no seat left and $e_{s3,A} > e_{s4,A}$ (because s4 rank A second).

- Conclusion: s2 has the same ranking of A, but he gets different eligibility score, based on other agent's ranking.

[AS16] prove that most mechanisms fall into the class of Report-Specif Priority and Cutoff Mechanisms, including BM and DA. Therefore, we could use the cutoff approach to consistently estimating assignment probabilities under BM and DA. However, [AS16] do not provide a detailed instruction of how to do it and their method of assigning eligibility score is relatively restrictive to make it easier for theoretic proof. In the following part, I provide a more practical way to assign eligibility score and explain how to estimate assignment probabilities with cutoffs under different mechanisms .

[AS16] define the function f as a mapping from reported rankings and lotteries to an eligibility score on $[0,1]$ for each school. They also define lotteries in a way that combines tie-breaking lotteries and priority types. The students with priority would receive a more favorable distribution of lotteries than students with no priority. For example, if one student has the highest priority in school A, second highest priority in school B, and no priority in school C, then his lotteries in school A, B, and C could be drawn from $(0.7,1]$, $(0.4,0.7]$, and $[0,0.4]$ respectively. This student's eligibility score for the three schools would also be mapped on $[0,1]$, no matter how he ranks them. This might be more convenient for theoretical proof, but it's not very convenient in calculating assignment probabilities. Especially when there are complicated priority types and many schools, it would be very difficult to assign lotteries and eligibility score using this way. In fact, the idea of eligibility scores and cutoffs has been widely used for assignment and experiment in practice. The general idea is to assign a large numbers to dominating factors and smaller numbers to less important factors in the mechanism, such that the biggest number of the lower priority group is smaller than the smallest number of the higher priority group. To make it easier for assignment probability calculation, I keep the tie-breaking lotteries on $[0,1]$. The priorities and ranking decide the integer part, following the assignment algorithm. As long as eligibility scores are assigned following this rule, it following the same results as [AS16], no matter how big the range is. The eligibility score I construct is in fact a monotonic transformation of the eligibility score in [AS16]. To better illustrate the construction of eligibility score, I will show one way to construct under BM and DA. There are many other ways to construct eligibility score based on the structure of priorities.

Remark *eligibility score under BM* Students are allowed to submit a ranking up to K schools. There are P different priorities. Students get a tie-breaking lottery for each school. Then the eligibility score for the rth ranking school s of student i is

$$(K-r)*(2+\sum_{p=1}^{P}p)+\sum_{p=1}^{P}(p+1)*\{\mathbf{P}_{isp}=1\}+L_{is}.$$

Remark *eligibility score under BA* Students are allowed to submit a ranking up to K schools. There are P different priorities. Students get a tie-breaking lottery for each school. Then the eligibility score for student i 's the rth ranking school, s, is

$$\sum_{p=1}^{P}(p+1)*\{\mathbf{P}_{isp}=1\}*S+L_{is},$$

where the $\mathbf{P}_{isp}$ is the indicator of whether student i has the pth LOWEST priority in school s. For the priority part, add 2 points if student i has the lowest priority type in school s, and add another 3 points if he has the second lowest priority type. So on so forth, add$P+1$ points if student i has the highest priority type in school s. The reason to add $p+1$ points to the pth priority type, rather than p points, is to distinguish the case where a student gets a tie-breaking lottery of 1 but has no priority and the case where a student who gets 0 as lottery but gets the lowest priority. If the priority is not addable, just remove the sum sign from the form and only count the highest priority. L_{is} is the tie-breaking lottery for student i in school s, which is drawn from $[0,1]$. The key ideal is that the upper criterion, either ranking or priority, is assigned a higher score than the lower criterion.

1.3 Simulations for cutoffs

In this section, I simulate assignments under different mechanisms using cutoffs. Following the random utility model, I define the utility of student i being assigned to school s as:

$$u_{i,s}=sibling_{i,s}*\beta_s+rating_s*\gamma-distance_{i,s}+\varepsilon_{is}$$

$$u_{i,0}=0$$

where the $sibling_{i,s}$ is an indicator of whether student i has sibling in school s; $rating_s$ is the school rating of school s; and $distance_{i,s}$ is the distance from the house of student i to school s. The outside option, $u_{i,0}$, is normalized to 0. The effect of home-school distance is normalized to -1.

Students can report up to 4 schools out of 4 options. Each school has the same capacity. Sibling status is randomly drawn from a binomial distribution. Additionally, one student may have siblings in multiple schools. School ratings for the 4 schools are 10, 8, 6, and 4. β measures the influence of siblings and I set it to [2, 2, 2, 2]. In practice, β is usually different across different schools. The reason I set it the same here is that the school rating would have relatively bigger influence than sibling status on preference of schools. γ could be taken from three number: 0.2, 1, and 2. Because school ratings are the same across different students, γ measures the correlation of preferences. The higher γ is, the higher is the correlation. Although error terms could be correlated, I use standard multivariate normal distribution for simplicity. The tie-breaking lottery for each student is randomly drawn from $[0,1]$ and keeps the same across different schools. In reality, the tie-breaking lottery could be different for different schools, whose effects will be examined in Chapter 2.

I only introduce one priority type, district student, to the this practice. Students have priority for their district school, which is sometimes called as walk-zone priority. The district school is defined as the school nearest to the student's home. Each student only has one district school. I conduct simulations both with and without district school priority. Different from theoretical papers, I do not define tie-breaking lottery as priority here. There are three information environments for simulation: (1) agents know their lotteries and the all information about cutoffs; (2) agents only know their lotteries and integer part of cutoffs, but do not know the decimal part of cutoffs; and (3) agents know both parts of cutoffs, but do not know their lotteries. The first environment would be ideal, because students have all the information they need to make the best response. However, the later two environments are closer to reality. In most school districts, students do not know their tie-breaking lotteries, but they know exactly the ranking and priority needed to have a chance of being assigned to a school, as well as the lottery needed. In some school districts, students know their tie-breaking lotteries and how to rank a school in order to get a chance of assignment, while they do not know exactly what lottery is needed to be assigned there.

Under all the simulation environments, I assume all agents are sophisticated and strategize optimally. More specifically, they submit a ranking that maximizes their expect utility. When there are multiple optimal solutions, that is, rankings that may yield to the same highest expected utility, I assume agents submit a ranking that's closest to their true preference. I compare the equilibria, cutoffs and assignment probabilities under these three information environments.

1.3.1 Cutoffs

Cutoffs consist of two parts: (1) the integer part which shows where they rank the school and the priority type, and (2) the decimal part which shows the lottery of the student who has the lowest eligibility score among students assigned to this school. If the cutoff is zero for a school, then this school is not filled up and still has seats left. It is possible, because I set the total capacity of the four school to be the same as the number of students. The utilities could be negative and students won't list these schools on the ranking in that case. If a student with negative utilities for certain schools is rejected by all the schools he ranked, then there will be schools not filled up and the cutoffs will be zero.

1.3.1.1 BM

The eligibility score under BM for student i's the rth ranking school, s, is $(4-r)*3+(\mathbf{P}_{is}=1)*2+L_i$, where $\mathbf{P}_{is}$ indicates whether school s is the district school for student i, and L_i is the tie-breaking lottery of student i. When there is no priority built in assignment, the integer part could be 9, 6, 3, and 0, each of which indicates students rank the school first, second, third, and fourth, respectively. When there is district priority built in assignment, each number of 9, 6, 3, and 0 indicates students rank a school first, second, third, and fourth respectively, and they do not have district priority in this school. Additionally, the integer part could also be 11, 8, 5, and 2, meaning students rank a school first, second, third, and fourth respectively and they have the district priority at this school. The decimal part indicates the lottery needed to be assigned to a school if the integer part of the eligibility sore is the same as the integer part of the school's cutoff.

Students first report their true preferences and get the assignment results and cutoffs. Then students update their rankings 1 by 1, based on the current cutoffs and update cutoffs after each student updating their ranking. Keep updating until the assignment does not change anymore. We reach the equilibrium. Starting with all truthful reports and starting with all manipulated reports (ranking the school where they have district priority first) lead to the same equilibrium. In all the cases, there's a unique assignment equilibrium. The cutoffs at equilibrium under different environment are shown in the Table 1.1.

1.3.1.2 DA

The eligibility score under DA for student i's the rth ranking school, s, is $(\mathbf{P}_{is}=1)*2+L_i$. In the case without district priority in assignment, the integer part of the cutoffs could only be 0, because how students rank schools does not count part of assignment criterion under DA. In the case with district priority in assignment, the integer part of the cutoffs could 0 and 2, where 2 means students have priority in the school.

Under unconstrained DA, where the number of schools allowed to rank is equal to total number of schools, it's strategy proof. However, it's not strategy proof under constrained DA, where the number of schools allowed to rank is smaller than total number of schools. Because students behave differently under constrained and unconstrained DA, I simulate both mechanisms. For constrained DA, I allow students to rank up to 3 schools out of 4 total schools. There's a unique equilibrium under both constrained and unconstrained DA. The cutoffs are shown in the Table 1.2. Panel a shows the cutoffs under unconstrained DA and panel b shows the cutoffs under constrained DA. We notice the results of constrained DA under all the different environments are very similar to unconstrained DA. It's partially because the constraint on the list is not sever here. If students are only allowed to submit 1 school, the results will be more different. The other reason is manipulation is very trivial under constrained DA, comparing to BM.

1.3.2 Manipulation

In this part, I demonstrate the manipulation behaviors under BM and constrained DA, with the three information environments. This helps us understand the manipulation in practice.

1.3.2.1 BM

Because some students only rank 2 schools out of 4, I only list the manipulation for the first two ranked schools. Those who manipulate their rankings usually start the manipulation from the first two schools. Therefore, showing the first two schools would be enough to show most manipulations. Tables 1.3 to 1.9 show the ratios of students with different manipulations. The denominators represent the numbers of students who have the true preferences as listed on the left of the table. The numerators show the numbers of students who rank schools as the top of the table. First, I show the manipulation under BM and where students know both parts of cutoffs and the lotteries. The results are shown in Table 1.3.

If only looking at the first ranked school versus the first true preference in the Table 1.3, it might seem that the environment without priority has slightly higher truth telling rate than the environment with priority. However, if we look at the second ranked school versus second truly preferred school, it's the opposite. It is also the case in Table 1.5, where students only know their lotteries and the integer part of cutoffs, but do not know the decimal part of cutoffs. It's because I set students to have priority at the closest school. At the same time, distance enters the utility of school and students have a very high utility at the closest school. As a result, for many students, the school where they have district priority is the same as their truly most preferred school. However, as priority becomes less important for the second ranked school, we observe the truth-telling rate decreases, and it is even lower in the scenario with priority built in mechanisms. For the same reason, if I use a different priority criterion that has a smaller influence on utility, such as a priority for students who live in a neighborhood with certain characteristics (e.g., high average test scores), then the truth-telling rate would be much lower even for the first ranked school. Therefore, it might be less problematic to treat reported first ranking as true first ranking under such information environment and if priority types are similar to district school priority that are highly correlated to the preference. In general, it's still problematic to treat reported rankings as true preferences, especially when agents do not know their lotteries. From Table 1.4, we can see that when agents do not know their lotteries, the rate of truth-telling is lower when assignment priorities are highly correlated with students' utilities and higher when there is no deterministic priority component. Table 1.4 and Table 1.5 are both manipulations under BM with partial information. Table 1.4 is under the environment where students only know the cutoffs, but do not know their lotteries. Table 1.5 is under the environment where students only know their lotteries and the integer part of cutoffs, but do not know the decimal part of cutoffs.

From Table 1.4 and Table 1.5, we notice that the rate of truth-telling is higher with partial information than with complete information about cutoffs and tie-breaking lotteries, because agents would rank a competitive school first as long as they believe their assignment probability is high enough. With complete information about cutoffs and tie-breaking lotteries, agents would avoid competitive schools because they know for sure that they cannot get in. The simulations with higher correlation among preferences of schools has lower truth-telling rate under all the environments. Higher correlation means popular schools are even more competitive, therefore more students avoid competitive schools.

1.3.2.2 DA

Because unconstrained DA is strategyproof, I only show the manipulations with constrained DA under the three environments. Results are shown in Table 1.6 to Table 1.8.

From Table 1.6 to Table 1.8, we can conclude that the manipulation under constrained DA is very small. There's even no manipulations when there's priority in assignment with complete information about cutoffs and tie-breaking lotteries, as shown in the right panel of Table 1.6. However, there are still manipulations in all the other cases, and the basic pattern of truth-telling rates under different scenarios is the same as under BM. With complete information (about cutoffs and tie-breaking lotteries) or when students do not know the decimal part of cutoffs, the rate of truth-telling is higher when assignment priorities are highly correlated with students' utilities. Truth-telling is lower when there is no deterministic priority component or when assignment priorities involve components that do not directly enter students' utilities. When agents do not know their lottery numbers, the opposite pattern holds (i.e., the rate of truth-telling is lower when assignment priorities are highly correlated with students' utilities, and it's higher when there is no deterministic priority component). There is more truth telling with partial information than with complete information about cutoffs and tie-breaking lotteries. With higher correlation among preferences of schools, the rate of truth-telling is lower under all the information environments.

1.3.3 Probability of assignment : cutoffs vs counting assignment

In this part, I compare the method of estimating assignment probabilities using the cutoff approach to the approach of counting assignments.

Under the cutoff approach, it is easy to calculate assignment probabilities if agents have complete information about cutoffs and tie-breaking lotteries. If student i is being considered for school s and his eligibility score at school s is higher than the cutoff of school s, then the assignment probability of him being assigned to school s is 1. If student i's eligibility score at school s is lower than the cutoff of school s, then the assignment probability of student i being assigned to school s is 0. With partial information about cutoffs or tie-breaking lotteries, it's more complicated.

If the student has not been rejected by any school, the following shows how to calculate assignment probabilities by using cutoffs, when students know cutoffs but do not know their tie-breaking lotteries:

- If the integer part of student i's eligibility score at school s is higher than the cutoff of school s, then the assignment probability of student i at school s is 1. Students do not know the decimal part of their eligibility score, because they do not know their tie-breaking lotteries.
- If the integer part of student i's eligibility score at school s is equal to the integer part of cutoff of school s, then the assignment probability of student i at school s is (1− decimal part of cutoff of school s).
- If the integer part of student i's eligibility score at school s i slower than the integer part of cutoff of school s, then the assignment probability of student i at school s is 0.

If the student has not been rejected by any school, the following shows how to calculate assignment probabilities by using cutoffs, when students know their tie-breaking lotteries but do not know the decimal part of cutoffs:

- If the integer part of student i's eligibility score at school s is higher than the integer part of cutoff of school s, then the assignment probability of student i at school s is 1. Students do not know the decimal part of cutoff of school s.
- If the integer part of student i's eligibility score at school s is equal to the integer part of cutoff of school s, then the assignment probability of student i at school s is equal to the tie-breaking lottery of student i at school s.
- If the integer part of student i's eligibility score at school s i slower than the integer part of cutoff of school s, then the assignment probability of student i at school s is 0.

The assignment probabilities calculated using instruction above is the unconditional probability when the student i is considered for school s. The assignment probability of student i at school s should be conditional on the probability that student i is rejected by all the other schools that are ranked before s. In other words, the assignment probabilities should be the probability of being rejected by all the schools that are ranked before s times the unconditional probability calculated as instructed above.

The following steps show how to calculate probabilities by counting assignment:

1. For a submitted ranking $R_1 = (C^1, C^2, \ldots, C^K)$, count the number of students who submit this ranking and have priority type p, denoted as $S_{R_1,p}$.

2. Among students who submit R_1, count how many students were assigned to each school on the list, i.e. $C^1, C^2, \ldots, C^K$. Denote the number of students assigned to $C^1, C^2, \ldots, C^K$ as $s_{R_1,p,C^1}, s_{R_1,p,C^2}, \ldots, s_{R_1,p,C^K}$.

3. Then the assignment probabilities of students with priority type p and who submit R_1 at schools $C^1, C^2, \ldots, C^K$ are $s_{R_1,p,C^1}/S_{R_1,p}$, $s_{R_1,p,C^2}/S_{R_1,p}, \ldots, s_{R_1,p,C^K}/S_{R_1,p}$.

4. Repeat step 1 to step 3 for each priority type.

5. Repeat step 1 to step 4 for each kind of submitted ranking.

As we could see it would be very complicated to calculate assignment probabilities by counting assignments. First, there could be many kinds of submitted rankings. Second, there could be many priority types. For example, there are eleven layers of priorities in Wake County Public School System (WCPSS). Students could have up to four out of the eleven priorities, which means there could be hundreds of different priority types. It requires calculation for each ranking and each priority type. On the other hand, we can use the matrix calculation to get the assignment probabilities easily with cutoffs.

When calculating the assignment probabilities in the scenario with priority built in mechanism, it needs to calculate the probabilities for both those who have priority and those who do not have priority. So, I use the scenario where there's no priority built in mechanism. I use the environment under BM with high correlation and where students do not know lottery but know the full cutoffs. To show the advantages of the cutoff approach, I calculate assignment probabilities with three different sample sizes, 1,000, 10,000, and 100,000. Results are shown in Table 1.9 to Table 1.11 respectively. As indicated in the tables, there are many feasible reports not observed under BM as a result of manipulation. To better understand the difference of assignment probability estimation using the cutoffs and by counting assignments, I also show the probabilities estimated under unconstrained DA with low correlation and no priority built in mechanism. Under unconstrained DA, people would report truthfully. With low correlation, school preferences are more diverse and therefore there is less competitiveness. As a result, more feasible reports are observed from submitted rankings. Because students would report truthfully under all the environments with unconstrained DA, the knowledge of lotteries or cutoffs does not matter here. Assignment probability estimations are shown in the Table 1.12 to 1.14, with a sample size of 1,000, 10,000 and 100,000 respectively .

From Table 1.9 to 1.14, we can conclude that there are much fewer feasible rankings observed in the reports under manipulable mechanisms. Even if all feasible rankings are submitted, they may not be reported frequently enough to get consistent probabilities. When the sample size is 1,000 or 10,000, the assignment probabilities estimated by counting assignment are way off the consistent estimates. Although when the sample size is very small, lotteries may not be evenly distributed on the $[0,1]$ interval and the assignment probabilities might be a little off, the probability estimation is still much better than the results from counting assignment. 1,000 is relatively big enough to get a set of lottery that is close to uniform distribution. What's more, such bias could be easily eliminated by bootstrapping [AS16], which is used in Chapter 2. As sample size increases, the assignment probabilities estimated by counting assignment gets closer to the results by cutoffs. However, such big sample is rare in reality. Moreover, as the number of schools grows, the counting assignment approach requires much larger sample size, because there are many more feasible rankings and each ranking needs to be reported frequently enough. On the other side, the cutoff approach only needs a sample that's sufficiently large to get a set of uniformly distributed lotteries. Therefore, the requirement of sample size is much smaller for the cutoff approach and it does not grow as the number of schools grows.

1.4 Conclusion

From the simulation results, we can see that manipulations are common with the Boston Mechanism (BM). There is a higher degree of manipulations when students' preferences over schools are highly correlated. This is common in practice: school districts tends to have some highly popular and some highly unpopular schools. An important conclusion is that the cutoff approach to calculating the assignment probabilities has many practical advantages for empirical work. First, it works better with a relatively small sample. Second, the sample size is not required to grow with number of schools. Third, the estimation approach does not require that we observe all feasible rankings in the data, which is required by the common approaches in the literature. Fourth, the computation is much easier when there is a complicated priority structure. Therefore, the cutoff approach works better in practice, especially under manipulable mechanisms or in school districts with many schools.

Table 1.1 Cutoffs under BM

	without priority				with priority			
High correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	9.7346	9.2176	3.6135	0.0038	9.9734	9.3341	5.4116	0.0099
No lottery	9.6922	6.5941	3.3110	0.00006	9.9572	8.8976	3.6925	0.0044
No decimal part	9.7346	6.7079	3.4593	0.00006	9.9734	8.5153	3.8557	0.0044
Mid correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	9.6857	9.2267	3.6153	0.00006	9.9150	9.3409	5.7441	0.00006
No lottery	9.6400	9.0260	3.3856	0	9.9088	9.2037	3.8858	0
No decimal part	9.6809	9.0255	3.4835	0	9.9150	9.0001	3.8976	0
Low correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	9.3531	9.0960	6.0391	0	9.5733	9.1718	6.2697	0
No lottery	9.3506	9.0822	3.2021	0	9.5712	9.1120	3.5329	0
No decimal part	9.3495	9.0875	3.1922	0	9.5750	9.0822	3.5401	0

Table 1.2 Cutoffs under DA

Panel a. Unconstrained DA

	without priority				with priority			
	A	**B**	**C**	**D**	**A**	**B**	**C**	**D**
High correlation	0.7346	0.4811	0.2572	0.00006	0.9734	0.9292	2.0092	0.00006
Mid correlation	0.6857	0.4424	0.2146	0	0.9388	0.9111	0.8207	0.0099
Low correlation	0.3531	0.2289	0.0845	0	0.6106	0.4010	0.2697	0

Panel b. Constrained DA

	without priority				with priority			
High correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	0.7346	0.4811	0.2572	0	0.9734	0.9292	2.0092	0
No lottery	0.7346	0.4841	0.2526	0	0.9731	0.9292	2.0092	0
No decimal part	0.7346	0.4841	0.1922	0	0.9734	0.9292	2.0092	0
Mid correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	0.6857	0.4424	0.2146	0	0.9388	0.9111	0.8207	0
No lottery	0.6857	0.4517	0.2341	0	0.9498	0.9239	0.8835	0
No decimal part	0.6857	0.4525	0.2091	0	0.9498	0.9239	0.8843	0
Low correlation	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Know both	0.3531	0.2289	0.0845	0	0.6106	0.4010	0.2697	0
No lottery	0.3531	0.2388	0.0845	0	0.6321	0.4667	0.2146	0
No decimal part	0.3531	0.2388	0.0834	0	0.6321	0.4667	0.2146	0

Table 1.3 Manipulation under BM, know both lottery and cutoffs

Panel a. High correlation $\gamma = 2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	590/975	343/975	42/975	0/975	596/975	343/975	36/975	0/975
true	B	0/25	25/25	0/25	0/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	9/25	0/25	15/25	1/25	9/25	0/25	15/25	1/25
true	B	0/946	512/946	383/946	51/946	0/946	478/946	424/946	44/946
	C	0/29	5/29	16/29	8/29	0/29	6/29	14/29	9/29

Panel b. Mid correlation $\gamma = 1$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	541/817	192/817	73/817	11/817	549/817	190/817	70/817	8/817
true	B	0/163	160/163	2/163	1/163	0/163	160/163	2/163	1/163
	C	0/20	0/20	20/20	0/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	60/165	2/165	76/165	27/165	60/165	2/165	76/165	27/165
true	B	0/625	368/625	195/625	62/625	0/625	345/625	215/625	65/625
	C	0/181	11/181	125/181	45/181	0/181	18/181	125/181	38/181
	D	0/29	2/29	6/29	21/29	0/29	2/29	4/29	23/29

Panel c. Low correlation $\gamma = 0.2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	371/408	13/408	16/408	8/408	371/408	13/408	14/408	10/408
true	B	0/267	265/267	0/267	2/267	0/267	265/267	0/267	2/267
	C	0/196	0/196	196/196	0/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	172/265	17/265	29/265	37/265	172/265	16/265	29/265	38/265
true	B	0/277	249/277	17/277	9/277	0/277	245/277	19/277	10/277
	C	0/241	4/241	226/241	7/241	0/241	4/241	227/241	6/241
	D	0/209	3/209	5/209	200/209	0/209	5/209	5/209	198/209

Table 1.4 Manipulation under BM, do not know lottery

Panel a. High correlation $\gamma = 2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	824/975	151/975	0/975	0/975	764/975	211/975	0/975	0/975
true	B	0/25	25/25	0/25	0/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	25/25	0/25	0/25	0/25	25/25	0/25	0/25	0/25
true	B	151/946	781/946	14/946	0/946	211/946	734/946	1/946	0/946
	C	0/29	7/29	22/29	0/29	0/29	12/29	17/29	0/29

Panel b. Mid correlation $\gamma = 1$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	722/817	95/817	0/817	0/817	689/817	128/817	0/817	0/817
true	B	0/163	163/163	0/163	0/163	1/163	162/163	0/163	0/163
	C	0/20	0/20	20/20	0/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	163/165	0/165	2/165	0/165	162/165	1/165	2/165	0/165
true	B	96/625	499/625	30/625	0/625	123/625	484/625	18/625	0/625
	C	9/181	48/181	123/181	0/181	9/181	92/181	79/181	0/181
	D	3/29	9/29	0/29	17/29	3/29	19/29	0/29	7/29

Panel c. Low correlation $\gamma = 0.2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	404/408	4/408	0/408	0/408	405/408	3/408	0/408	0/408
true	B	0/267	267/267	0/267	0/267	0/267	267/267	0/267	0/267
	C	0/196	0/196	196/196	0/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	126/265	0/265	3/265	0/265	127/265	0/265	2/265	0/265
true	B	44/277	105/277	6/277	0/277	44/277	96/277	2/277	0/277
	C	68/241	38/241	55/241	0/241	70/241	42/241	35/241	0/241
	D	69/209	27/209	0/209	35/209	69/209	35/209	0/209	22/209

Table 1.5 Manipulation under BM, do not know decimal part of cutoffs

Panel a. High correlation $\gamma = 2$

		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	780/975	195/975	0/975	0/975	851/975	124/975	0/975	0/975
true	B	0/25	25/25	0/25	0/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	25/25	0/25	0/25	0/25	25/25	0/25	0/25	0/25
true	B	195/946	722/946	29/946	0/946	124/946	798/946	24/946	0/946
	C	0/29	7/29	22/29	0/29	0/29	11/29	18/29	0/29

Panel b. Mid correlation $\gamma = 1$

		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	718/817	99/817	0/817	0/817	730/817	87/817	0/817	0/817
true	B	0/163	163/163	0/163	0/163	0/163	163/163	0/163	0/163
	C	0/20	0/20	20/20	0/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	156/165	0/165	9/165	0/165	165/165	0/165	0/165	0/165
true	B	104/625	455/625	65/625	0/625	93/625	508/625	24/625	0/625
	C	9/181	48/181	123/181	0/181	9/181	92/181	79/181	0/181
	D	3/29	12/29	0/29	14/29	3/29	19/29	0/29	7/29

Panel c. Low correlation $\gamma = 0.2$

		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	401/408	7/408	0/408	0/408	407/408	1/408	0/408	0/408
true	B	0/267	267/267	0/267	0/267	1/267	266/267	0/267	0/267
	C	0/196	0/196	196/196	0/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	126/265	0/265	4/265	0/265	126/265	0/265	2/265	0/265
true	B	46/277	101/277	6/277	0/277	42/277	98/277	2/277	0/277
	C	68/241	40/241	53/241	0/241	70/241	42/241	35/241	0/241
	D	69/209	28/209	0/209	34/209	69/209	35/209	0/209	22/209

Table 1.6 Manipulation under constrained DA, know both lottery and cutoffs

Panel a. High correlation $\gamma = 2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	590/975	225/975	119/975	41/975	975/975	0/975	0/975	0/975
true	B	0/25	21/25	3/25	1/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	13/25	0/25	9/25	3/25	25/25	0/25	0/25	0/25
true	B	0/946	574/946	219/946	120/946	0/946	946/946	0/946	0/946
	C	0/29	9/29	16/29	3/29	0/29	0/29	29/29	0/29

Panel b. Mid correlation $\gamma = 1$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	541/817	135/817	100/817	41/817	817/817	0/817	0/817	0/817
true	B	0/163	149/163	8/163	6/163	0/163	163/163	0/163	0/163
	C	0/20	0/20	19/20	1/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	118/165	2/165	25/165	16/165	165/165	0/165	0/165	0/165
true	B	0/625	412/625	124/625	71/625	0/625	625/625	0/625	0/625
	C	0/181	35/181	125/181	19/181	0/181	0/181	181/181	0/181
	D	0/29	4/29	3/29	20/29	0/29	0/29	0/29	29/29

Panel c. Low correlation $\gamma = 0.2$

		Without Priority				With Priority			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	371/408	11/408	18/408	8/408	408/408	0/408	0/408	0/408
true	B	0/267	261/267	1/267	5/267	0/267	267/267	0/267	0/267
	C	0/196	0/196	194/196	2/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	248/265	4/265	5/265	4/265	265/265	0/265	0/265	0/265
true	B	0/277	259/277	10/277	5/277	0/277	277/277	0/277	0/277
	C	0/241	5/241	225/241	6/241	0/241	0/241	241/241	0/241
	D	0/209	4/209	2/209	199/209	0/209	0/209	0/209	209/209

Table 1.7 Manipulation under constrained DA, do not know lottery

Panel a. High correlation $\gamma = 2$									
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	975/975	0/975	0/975	0/975	953/975	22/975	0/975	0/975
true	B	0/25	25/25	0/25	0/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	25/25	0/25	0/25	0/25	24/25	0/25	1/25	0/25
true	B	0/946	943/946	3/946	0/946	0/946	921/946	25/946	0/946
	C	0/29	7/29	22/29	0/29	0/29	9/29	20/29	0/29

Panel b. Mid correlation $\gamma = 1$									
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	816/817	1/817	0/817	0/817	809/817	8/817	0/817	0/817
true	B	0/163	163/163	0/163	0/163	0/163	163/163	0/163	0/163
	C	0/20	0/20	20/20	0/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	158/165	3/165	4/165	0/165	165/165	0/165	0/165	0/165
true	B	5/625	616/625	4/625	0/625	6/625	602/625	17/625	0/625
	C	6/181	48/181	126/181	0/181	6/181	77/181	97/181	0/181
	D	2/29	9/29	0/29	18/29	3/29	16/29	0/29	10/29

Panel c. Low correlation $\gamma = 0.2$									
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	408/408	0/408	0/408	0/408	408/408	0/408	0/408	0/408
true	B	0/267	267/267	0/267	0/267	0/267	267/267	0/267	0/267
	C	0/196	0/196	196/196	0/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	142/265	4/265	5/265	4/265	138/265	0/265	0/265	0/265
true	B	41/277	116/277	0/277	0/277	41/277	102/277	0/277	0/277
	C	59/241	37/241	69/241	0/241	67/241	40/241	42/241	0/241
	D	53/209	21/209	0/209	65/209	57/209	31/209	0/209	43/209

Table 1.8 Manipulation under trunacted DA, do not know decimal part of cutoffs

		Panel a. High correlation $\gamma = 2$							
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	975/975	0/975	0/975	0/975	975/975	0/975	0/975	0/975
true	B	0/25	25/25	0/25	0/25	0/25	25/25	0/25	0/25
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	25/25	0/25	0/25	0/25	25/25	0/25	0/25	0/25
true	B	0/946	946/946	0/946	0/946	0/946	946/946	0/946	0/946
	C	0/29	7/29	22/29	0/29	0/29	9/29	20/29	0/29
		Panel b. Mid correlation $\gamma = 1$							
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	817/817	0/817	0/817	0/817	817/817	0/817	0/817	0/817
true	B	0/163	163/163	0/163	0/163	0/163	163/163	0/163	0/163
	C	0/20	0/20	20/20	0/20	0/20	0/20	20/20	0/20
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	165/165	0/165	0/165	0/165	165/165	0/165	0/165	0/165
true	B	5/625	620/625	0/625	0/625	6/625	619/625	0/625	0/625
	C	6/181	48/181	126/181	0/181	6/181	77/181	97/181	0/181
	D	2/29	9/29	0/29	18/29	3/29	16/29	0/29	10/29
		Panel c. Low correlation $\gamma = 0.2$							
		Without Priority				**With Priority**			
		1st ranked school				1st ranked school			
		A	B	C	D	A	B	C	D
1st	A	408/408	0/408	0/408	0/408	408/408	0/408	0/408	0/408
true	B	0/267	267/267	0/267	0/267	0/267	267/267	0/267	0/267
	C	0/196	0/196	196/196	0/196	0/196	0/196	196/196	0/196
	D	0/129	0/129	0/129	129/129	0/129	0/129	0/129	129/129
		2nd ranked school				2nd ranked school			
		A	B	C	D	A	B	C	D
2nd	A	142/265	0/265	0/265	0/265	138/265	0/265	0/265	0/265
true	B	41/277	116/277	0/277	0/277	41/277	102/277	0/277	0/277
	C	59/241	37/241	69/241	0/241	67/241	40/241	42/241	0/241
	D	53/209	21/209	0/209	65/209	57/209	31/209	0/209	43/209

Table 1.9 Assignment Probabilities under BM, do not know lottery, high correlation, N=1,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.3078	0.2809	0.2834	0.1279	0.3077	0.0923	0.2859	0.3141
ABDC	0.3078	0.2809	0	0.4112	0.1250	0.2500	0	0.6250
ACBD	0.3078	0	0.6922	0	0.2500	0	0.7500	0
ACDB	0.3078	0	0.6922	0	.	.	.	.
ADBC	0.3078	0	0	0.6922	.	.	.	.
ADCB	0.3078	0	0	0.6922	.	.	.	.
BACD	0	1	0	0	0	1	0	0
BADC	0	1	0	0	.	.	.	.
BCAD	0	1	0	0	.	.	.	.
BCDA	0	1	0	0	.	.	.	.
BDAC	0	1	0	0	.	.	.	.
BDCA	0	1	0	0	.	.	.	.
CABD	0	0	1	0	.	.	.	.
CADB	0	0	1	0	.	.	.	.
CBAD	0	0	1	0	.	.	.	.
CBDA	0	0	1	0	.	.	.	.
CDAB	0	0	1	0	.	.	.	.
CDBA	0	0	1	0	.	.	.	.
DABC	0	0	0	1	.	.	.	.
DACB	0	0	0	1	.	.	.	.
DBAC	0	0	0	1	.	.	.	.
DBCA	0	0	0	1	.	.	.	.
DCAB	0	0	0	1	.	.	.	.
DCBA	0	0	0	1	.	.	.	.

Notes: Cutoffs are 9.6922, 6.5941, 3.3110, and 0.00006. '.' means the ranking is not observed in the reports.

Table 1.10 Assignment Probabilities under BM, do not know lottery, high correlation, N=10,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.3112	0.2589	0.2854	0.1445	0.3111	0.0682	0.2961	0.3247
ABDC	0.3112	0.2589	0	0.4299	0.3662	0.0986	0	0.5352
ACBD	0.3112	0	0.6888	0	0.3108	0	0.6892	0
ACDB	0.3112	0	0.6888	0	.	.	.	.
ADBC	0.3112	0	0	0.6888	.	.	.	.
ADCB	0.3112	0	0	0.6888	.	.	.	.
BACD	0	1	0	0	0	1	0	0
BADC	0	1	0	0	.	.	.	.
BCAD	0	1	0	0	.	.	.	.
BCDA	0	1	0	0	.	.	.	.
BDAC	0	1	0	0	.	.	.	.
BDCA	0	1	0	0	.	.	.	.
CABD	0	0	1	0	.	.	.	.
CADB	0	0	1	0	.	.	.	.
CBAD	0	0	1	0	.	.	.	.
CBDA	0	0	1	0	.	.	.	.
CDAB	0	0	1	0	.	.	.	.
CDBA	0	0	1	0	.	.	.	.
DABC	0	0	0	1	.	.	.	.
DACB	0	0	0	1	.	.	.	.
DBAC	0	0	0	1	.	.	.	.
DBCA	0	0	0	1	.	.	.	.
DCAB	0	0	0	1	.	.	.	.
DCBA	0	0	0	1	.	.	.	.

Notes: Cutoffs are 9.6888, 6.6242, 3.3362, and 0.00006.

Table 1.11 Assignment Probabilities under BM, do not know lottery, high correlation, N=100,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.3089	0.2655	0.2890	0.1366	0.3084	0.0759	0.2947	0.3209
ABDC	0.3089	0.2655	0	0.4256	0.3054	0.0794	0	0.6152
ACBD	0.3089	0	0.6911	0	0.3258	0	0.6742	0
ACDB	0.3089	0	0.6911	0	.	.	.	.
ADBC	0.3089	0	0	0.6911	0.500	0	0	0.500
ADCB	0.3089	0	0	0.6911	.	.	.	.
BACD	0	1	0	0	0	1	0	0
BADC	0	1	0	0	.	.	.	.
BCAD	0	1	0	0	.	.	.	.
BCDA	0	1	0	0	.	.	.	.
BDAC	0	1	0	0	.	.	.	.
BDCA	0	1	0	0	.	.	.	.
CABD	0	0	1	0	0	0	1	0
CADB	0	0	1	0	.	.	.	.
CBAD	0	0	1	0	.	.	.	.
CBDA	0	0	1	0	.	.	.	.
CDAB	0	0	1	0	.	.	.	.
CDBA	0	0	1	0	.	.	.	.
DABC	0	0	0	1	.	.	.	.
DACB	0	0	0	1	.	.	.	.
DBAC	0	0	0	1	.	.	.	.
DBCA	0	0	0	1	.	.	.	.
DCAB	0	0	0	1	.	.	.	.
DCBA	0	0	0	1	.	.	.	.

Notes: Cutoffs are 9.6911, 6.6158, 3.3209, and 0.000000156.

Table 1.12 Assignment Probabilities under unconstrained DA, low correlation, N=1,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.6469	0.2723	0.074	0.0068	0.6591	0.1364	0.1591	0.0455
ABDC	0.6469	0.2723	0	0.0808	0.7429	0.0571	0	0.2
ACBD	0.6469	0.023	0.3233	0.0068	0.5333	0	0.3333	0.1333
ACDB	0.6469	0	0.3233	0.0299	0.6786	0	0.2857	0.0357
ADBC	0.6469	0	0	0.3531	0.5217	0	0	0.4783
ADCB	0.6469	0	0	0.3531	0.6	0	0	0.4
BACD	0.148	0.7711	0.074	0.0068	0	0.75	0.2143	0.0357
BADC	0.148	0.7711	0	0.0808	0	0.8929	0	0.1071
BCAD	0.0125	0.7711	0.2095	0.0068	0	0.8387	0.0968	0.0645
BCDA	0	0.7711	0.2095	0.0194	0	0.9474	0.0526	0
BDAC	0	0.7711	0	0.2289	0	0.5556	0	0.4444
BDCA	0	0.7711	0	0.2289	0	0.8571	0	0.1429
CABD	0.0547	0.023	0.9155	0.0068	0	0	0.8182	0.1818
CADB	0.0547	0	0.9155	0.0299	0	0	0.8947	0.1053
CBAD	0.0125	0.0652	0.9155	0.0068	0	0	1	0
CBDA	0	0.0652	0.9155	0.0194	0	0	1	0
CDAB	0	0	0.9155	0.0845	0	0	1	0
CDBA	0	0	0.9155	0.0845	0	0	0.9	0.1
DABC	0	0	0	1	0	0	0	1
DACB	0	0	0	1	0	0	0	1
DBAC	0	0	0	1	0	0	0	1
DBCA	0	0	0	1	0	0	0	1
DCAB	0	0	0	1	0	0	0	1
DCBA	0	0	0	1	0	0	0	1

Notes: Cutoffs are 0.3531, 0.2289, 0.0845, and 0.

Table 1.13 Assignment Probabilities under unconstrained DA, low correlation, N=10,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.6276	0.2905	0.077	0.0049	0.6702	0.1754	0.1204	0.034
ABDC	0.6276	0.2905	0	0.0819	0.6429	0.1463	0	0.2109
ACBD	0.6276	0.0174	0.3501	0.0049	0.6142	0	0.3234	0.0623
ACDB	0.6276	0	0.3501	0.0224	0.6644	0	0.2915	0.0441
ADBC	0.6276	0	0	0.3724	0.6602	0	0	0.3398
ADCB	0.6276	0	0	0.3724	0.6537	0	0	0.3463
BACD	0.138	0.78	0.077	0.0049	0	0.8239	0.1289	0.0472
BADC	0.138	0.78	0	0.0819	0	0.7509	0	0.2491
BCAD	0.0083	0.78	0.2068	0.0049	0	0.7418	0.1885	0.0697
BCDA	0	0.78	0.2068	0.0132	0	0.788	0.1739	0.038
BDAC	0	0.78	0	0.22	0	0.7713	0	0.2287
BDCA	0	0.78	0	0.22	0	0.7076	0	0.2924
CABD	0.0377	0.0174	0.94	0.0049	0	0	0.9592	0.0408
CADB	0.0377	0	0.94	0.0224	0	0	0.9107	0.0893
CBAD	0.0083	0.0468	0.94	0.0049	0	0	0.9082	0.0918
CBDA	0	0.046833	0.939959	0.013207	0	0	0.9231	0.0769
CDAB	0	0	0.939959	0.060041	0	0	0.9071	0.0929
CDBA	0	0	0.939959	0.060041	0	0	0.9138	0.0862
DABC	0	0	0	1	0	0	0	1
DACB	0	0	0	1	0	0	0	1
DBAC	0	0	0	1	0	0	0	1
DBCA	0	0	0	1	0	0	0	1
DCAB	0	0	0	1	0	0	0	1
DCBA	0	0	0	1	0	0	0	1

Notes: Cutoffs are 0.3724, 0.2200, 0.0600, and 0.

Table 1.14 Assignment Probabilities under unconstrained DA, low correlation, N=100,000

	Calculating by cutoffs				Calculating by assignment			
	A	B	C	D	A	B	C	D
ABCD	0.6464	0.2835	0.0653	0.0048	0.6516	0.1491	0.1285	0.0709
ABDC	0.6464	0.2835	0	0.0701	0.6558	0.1563	0	0.1879
ACBD	0.6464	0.0194	0.3294	0.0048	0.636	0	0.291	0.073
ACDB	0.6464	0	0.3294	0.0242	0.6382	0	0.2885	0.0734
ADBC	0.6464	0	0	0.3536	0.6414	0	0	0.3586
ADCB	0.6464	0	0	0.3536	0.6431	0	0	0.3569
BACD	0.1282	0.8017	0.0653	0.0048	0	0.8032	0.1302	0.0666
BADC	0.1282	0.8017	0	0.0701	0	0.7989	0	0.2011
BCAD	0.0088	0.8017	0.1847	0.0048	0	0.8042	0.1264	0.0694
BCDA	0	0.8017	0.1847	0.0136	0	0.802	0.123	0.075
BDAC	0	0.8017	0	0.1983	0	0.7861	0	0.2139
BDCA	0	0.8017	0	0.1983	0	0.7918	0	0.2082
CABD	0.0442	0.0194	0.9317	0.0048	0	0	0.9309	0.0691
CADB	0.0442	0	0.9317	0.0242	0	0	0.9277	0.0723
CBAD	0.0088	0.0548	0.9317	0.0048	0	0	0.9334	0.0666
CBDA	0	0.0548	0.9317	0.0136	0	0	0.9213	0.0787
CDAB	0	0	0.9317	0.0683	0	0	0.9497	0.0503
CDBA	0	0	0.9317	0.0683	0	0	0.9278	0.0722
DABC	0	0	0	1	0	0	0	1
DACB	0	0	0	1	0	0	0	1
DBAC	0	0	0	1	0	0	0	1
DBCA	0	0	0	1	0	0	0	1
DCAB	0	0	0	1	0	0	0	1
DCBA	0	0	0	1	0	0	0	1

Notes: Cutoffs are 0.3536, 0.1983, 0.0683, and 0.

CHAPTER

2

DESIGNING SCHOOL CHOICE MECHANISMS: A STRUCTURAL MODEL AND DEMAND ESTIMATION

2.1 Introduction

The importance of efficient and fair allocation of public school seats has led to collaborations between economists and school districts to improve outcomes associated with student assignment. I conduct a structural analysis of rich assignment data to provide insight into the design of school choice mechanisms. These mechanisms determine the procedure for conducting an assignment. Numerous design choices must be made by a school district as they implement an assignment mechanism and different design choices can make important differences in the resulting assignment. There have been many theoretical papers studying the tradeoff between efficiency and fairness under different mechanisms and how different design features would affect the assignment results.

This chapter studies several important design choices that school districts face with school choice. Most importantly, policymakers have to decide which mechanism to use. The vast majority of school districts use either BM or DA; TTC was used in New Orleans for one year before being replaced with DA [Abd17]. While the choice among mechanisms has received a great deal of attention among theoretical and experimental economists, the literature using data from school districts remains limited. Further, the choice among mechanisms is only one of several decisions. To implement a mechanism in practice, a school district must make several additional design choices to achieve its full set of objectives. Districts must decide how students will submit their preferences over schools, which includes features such as how many schools can be reported on students' submitted preferences. These more "mundane" design choices are nevertheless important in practice and, as my results show, can matter a lot in terms of important outcomes such as economic efficiency. Further, districts often choose to reserve some seats to achieve affirmative action goals that result in preferential treatment of students with free or reduced price lunch status or students from certain demographic groups.[1] In other cases, districts reserve seats to achieve academic goals that result in preferential treatment of high-achieving students. My focus on the real world design questions in school choice is in the spirit of [Pat16], who uses the practical experiences of economists who have collaborated with school districts to make recommendations for designing school choice mechanisms that work well in the field.

To study these and other design choices, I use data from the Wake County Public School System (WCPSS) in North Carolina.[2] WCPSS switched to DA in 2015 after using BM for many years. I analyze data from the academic year 2014-2015, which is the last year that BM was used. In 2014, WCPSS allowed students to list up to three schools, which is highly constrained because there are 23 elementary schools, 9 middle schools, and 4 high schools.[3] The constrained list places applicants in a highly competitive strategic setting, beyond the fact that BM is a manipulable mechanism. My structural model provides a framework for estimating the demand for schools under manipulable mechanisms. Given that BM remains the most used mechanism in practice and constrained lists are very common, demand estimation under manipulable mechanisms is particularity important to extend theoretical results to the real world in terms of welfare analysis. In addition, demand estimation for schools is also important to school districts and policymakers, as it demonstrates the popularity of schools and gives guidance for how to allocate resources. My work adds to the small but growing literature on structurally estimating manipulable mechanisms ([Aja13], [Fac15], [AS16], [He16], [Hwa16], [Cal16], [Abd15], [Haa16], [Nar16] and [Kap17]).[4]

My structural model builds on these related papers, following [AS16] most closely. As is done in the related literature, I assume students are rational agents who submit a preference list of schools in order to maximize their expected utility, given the probability of assignment to a given school when

[1] Reserved seats are common in primary and secondary education. For example, Cambridge, MA has reserved seats for students with free and reduced price lunch status [AS16].

[2] The Wake County Public School System is the 15th largest school system in the United States, with 159,549 students in 2016-2017 [WCP16]. This is an interesting setting to study because of the school district's size and its diversity in terms of race/ethnicity and urbanicity.

[3] The schools of interest in WCPSS are magnet schools. There are 177 schools in the system overall, the majority of which are non-choice-based schools. Magnet schools are partially-choice-based assignment schools aimed at "reduc[ing] high concentrations of poverty and support[ing] diverse populations." See `http://www.wcpss.net/magnet`.

[4] Although [Abd15], [Fac15], and [Nar16] study school demand under DA, the mechanisms are manipulable due to constrained lists.

listed at a given position. The estimation is Bayesian, which is also the approach in [AS16] and [Kap17]. The analysis first estimates the [AS16] model, where all students are expected utility maximizers and the effects of student-level covariates are estimated in the utility function, along with the demand for each school. Then, I add school-level characteristics to learn how observable features of a school matter in determining its demand. The results suggest that school characteristics explain most of the variation in utility across schools, while individual characteristics play a smaller role. Finally, I estimate a model that includes heterogeneity among students in terms of their strategic play. Specifically, some students are *sophisticated* agents, who maximize their expected utility according to unbiased beliefs, and other students are *sincere* agents, who submit their preferences truthfully irrespective of the utility maximizing report.[5]

While other related papers have estimated models with heterogeneous sophistication, my innovation is to exploit a unique aspect of my data in order to identify students' level of sophistication. I use data from the application website where students submitted their preferences. In 2014, WCPSS provided information about the demand of schools during the application period, specifically showing the number of applicants who had previously ranked a given school first on their preference list. A student can change her ranking as many times as she wishes and can log into the website multiple times to observe the growth in demand for each school. [Dur17a] argue that students who log in only once, and whose login was earlier than the final few days of the application period, are unlikely to strategically respond to the information about demand and thus are more likely to submit truthfully (sincere students). In contrast, students who repeatedly log in and whose final login is near the end of the application period are more likely to strategically respond to the demand shown on the website (sophisticated students). I use the frequency and timing of students' logins to proxy for their level of sophistication. My heterogeneous sophisticated model uses this information to identify students' type (sophisticated or sincere) and then estimates their true cardinal preferences over schools given the type.

I present results on several outcomes of interest. First, I analyze strategic behavior to understand what reported preferences students submit to maximize their expected utility. This allows me to understand the strategic incentives that a manipulable mechanism provides students and, in my heterogeneous sophistication model, which students respond to them (sophisticated) and which students do not (sincere). Second, economic efficiency is measured by aggregate utility associated with the assignment. Efficiency has been the focus of many papers on school choice mechanisms. It is well known that, among strategyproof mechanisms, TTC is efficient, while DA is not. BM cannot Pareto-dominate DA when preferences are submitted simultaneously [ES06], but can Pareto-dominate DA when preferences are submitted sequentially [Dur17b]. Third, fairness is defined in the matching literature as the elimination of justified envy, which removes all instances in which student

[5]The terms sophisticated and sincere are simply shorthand for whether true preferences are submitted. This terminology follows [PS08].

j is assigned to a school that is preferred by student i to i's assignment when i has higher priority than j at the school. While both are strategyproof mechanisms, DA is fair and TTC is efficient. I compare DA and TTC to BM in terms of fairness as well. Besides the choice of mechanisms, I also compare the efficiency and fairness for other factors in school choice design: length of report, tie-breaking rules, and reserve seats. [AN17] study the influence of tie-breaking rules and find opposite efficiency-fairness tradeoffs between tie-breaking rules in over-demanded and under-demanded markets. However, the real cases are usually a mix of these two, with some schools over-demanded and some under-demanded. I study how the tie-breaking rules affect the fairness and efficiency in a market with both over-demanded and under-demanded schools.

My results show that the design of a school choice mechanism has important implications for the outcomes associated with student assignment. We find that different mechanisms lead to very different outcomes in terms of efficiency and fairness. There is a small, but noticeable, efficiency loss of using DA rather than TTC, but TTC has a lot of justified envy. More importantly, consistent with [PS08], switching away from BM to a strategyproof mechanism such as DA levels the playing field by reducing the disadvantage faced by less strategically sophisticated students. My results are very consistent with predictions from matching theory.

I also study additional design choices that a school district must make when implementing a choice-based assignment procedure. In large assignment problems such as the one I study, it is common to have constraints on the number of schools that a student can report. I find that constrained lists do not affect efficiency much but lead to more justified envy. Further, the use of single versus multiple tie breaking does not affect efficiency or fairness to any meaningful degree. In contrast, the presence of reserved seats that favor certain groups of students has large effects. In this setting where some seats are reserved for academically gifted students, reserved seats result in less justified envy but conflicting results for efficiency.

The rest of this chapter is organized as follows: Section 2 summarizes Related literate, Section 3 introduces the magnet school assignment mechanism in WCPSS from the academic year 2014-2015, Section 4 provides the model for structural estimation, Section 5 estimates the demand of schools, Section 6 analyzes ranking behaviors, Section 7 presents the results of counterfactual analyses, and Section 8 concludes.

2.2 Related literate

Different papers have different assumptions about students' believes of the assignments. [Cal16] assumes students have enough information to anticipate the correct distribution of rankings that others report. In other words, they believe students know their probabilities of each possible ranking in the year of application. This assumption is strict, considering there's no evidence in their sample to support this assumption. To ease this concern, [AS16] have two sets of beliefs for estimation. The first one follows the same assumption that all or some students have the right realization of the distribution of rankings in the current year. In the other set, they assume students do not know the distribution of others' rankings in the current year. Instead, they assume students fully realize the distribution of ranking in the previous year and students submit their rankings responding to the assignment results in the previous year. The second set of belief is also problematic, as it assumes rational students would respond to the equilibrium of previous year. If only one student changes the ranking in response to an existing equilibrium, the equilibrium would not change. However, if all the students change their rankings at the same time in response to the existing equilibrium, then the equilibrium will change. Sophisticated students would understand this, re-assess the equilibrium, and respond to the new equilibrium, but not the previous equilibrium. In the samples of these two papers, neither of them has evidence showing that students have correct beliefs about distribution of rankings. [He16] first accepts the same assumption in his benchmark model and then expands it to a model where students could have wrong beliefs. He groups submitted rankings and derives subject belief range based on mistake types. His model is partially identified and therefore he uses moments inequalities for estimation. [Hwa16] assumes students strategize according to certain decision rules, rather than an exact belief of assignment. Therefore, he does not assume students have correct belief about distribution of rankings of all the others. Instead, he makes two assumptions: (1) given any two schools, student know which school is more competitive and (2) do not put a school on your ranking unless you prefer it to high-probability ones. These two assumptions are truly weaker than the assumption that students have correct belief of distribution of rankings of all the other students. But they are still hard to hold in reality without additional information to students. It might be easy to decide which school has higher probability or fills up sooner between a competitive school and a noncompetitive school. But it would still be very difficult to decide which is even more competitive between two competitive schools. In fact, if students could easily decide which is more competitive between two similarly competitive schools, then they know exactly the order of competitiveness of all the schools, which is only possible if students at know the distribution of rankings of all other students. Besides, his assumptions makes him unable to get exact assignment probabilities of students and hence he could only use moments inequalities, same as [He16]. Moments inequality can only generate bounds estimation. The estimated range could be very wide and non-informative. The assumption of full realization of distribution of rankings benefits me to get the exact assignment probabilities of each student for each feasible ranking, which allows me to get point estimation. [Kap17] take advantages of the full identification of probability but also believe students may have wrong beliefs. In their model, they identify the belief error for each student for each school at each ranking position. They first estimate the parameters of belief errors using data from a survey on beliefs from part of the applicants. Then they back up belief errors for all students and estimate the demand with individual belief errors. The additional survey data help them pin down subjective beliefs. Unfortunately, the survey is not available in most areas and is subject to reporting errors. I follow the assumption of full realization of distribution of rankings as in [AS16], [Cal16], and [Kap17]. WCPSS provided additional information that lets applicants realize the ranking distribution and also supports for this assumption. Students in WCPSS can see how many students have ranked each school first and they could update their rankings based on this information. With this information, they would easily know which school is more competitive this year. There were also many school information sections that gave parents information of previous application results. Combining numbers of students who rank each school first with the assignment results from the previous years, students could have much better assessment of assignment probabilities in the current year in WCPSS.

All the structural empirical papers use a random utility model and assume each student have a random utility for each school. Except [He16] and [Fac15] where the error term is assumed to follow type I extreme value distribution, all the other papers assume that the error term follows a multivariate distribution. All of them assume sophisticated students submit a ranking that maximizes their expected utility and naive students report their true preferences. As mentioned above, [He16] and [Hwa16] only provide partial identification and use moments inequalities for estimation. [AS16], [Cal16], and [Kap17] all have full identification and could get point estimates. Given the exact assignment probability for each student in each school for each feasible ranking, the submitted ranking of a rational student reveals his true preferences, because a rational student will only choose the ranking that yields to the highest expect utility. They assume the error term follows normal distribution, the utility model becomes a multivariate probit model.

It is well known that it's difficult to estimate multivariate probit model. One advantage of choosing logit model for discrete choice problem is the easiness in estimation. The reason why all the full identification papers choose to use probit model, but not logit, is that the errors are more likely to be correlated and heteroskedastic under the school choice context. It's widely believed that the personal characteristics play a large role in school choices and residential choices such as their work locations, income levels, and, races, and many of these individual characteristics may not be observed. As a result, the error terms are likely to be correlated and possibly heteroskedastic. Probit model, which is known to be difficult to estimate, is therefore chosen to handle the correlation and heteroskedasticity among choice options. Tradition maximum likelihood estimation is computationally intense for in this case. Alternative estimation methods are developed, including both classical and Bayesian traditions. On the classical side, some variations of traditional maximum likelihood estimation have been developed and the most widely used one is the maximum simulated likelihood estimate developed by [McF89], [Gew91], [Haj90], and [Kea94]. This approach relies on a Geweke, Hajivassiliou, and Keane (GHK) choice probability simulator. On the Bayesian side, [AC93], [MR94], and [AR99] develop an approach that uses Gibbs sampler with data augmentation to estimate multivariate probit model. [MR94] also use Hierarchical Bayesian model to estimate random coefficient multivariate probit model with panel data. [Bol97] compares the maximum simulated likelihood approach and the Gibbs sampler approach in multivariate probit estimation. They find the two approaches yield to similar estimation results, but the Gibbs sampler approach is much simpler to implement both conceptually and computationally. More importantly, for their particular setting, the Gibbs sampler approach is about twice faster as fast as the maximum simulated likelihood.

Several papers have tried to model the heterogeneity in sophistication levels in students. Although [AS16] introduce heterogeneous types of students to their model, their model is still experimental without many details. They start with assuming everyone is naive and update the type of each student in the Markov Chain Monte Carlo. For students who they cannot decide whether they are naive or sophisticated, they draw the type of this student from a binomial distribution. They have not provided any proof showing the estimates from this approach are consistent. Due to the weak identification, I find their approach is very sensitive to priors after applying it to my sample. [He16] and [Hwa16] randomly choose different percentages of their sample to be naive students and conduct welfare analysis under different percentages. They may wrongly choose a sophisticated student to be naive and vise versa. The estimates of utilities would be biased because the revealed preference is not true preference in these cases. Besides, they do not know what the approximate percentage of naive students is in the sample. [Cal16] uses a mixture model to identify the sophistication levels, naive or sophisticated, in the maximum likelihood estimation process. [He16] and [Kap17] do not focus on the sophistication levels of students. In these two papers, students are fully sophisticated but differ in beliefs. Students may make mistake because they have wrong beliefs about the distribution of rankings of others, or in other words, the assignment probabilities. The belief error approach may capture a continuum between the best response and the true response (naive), however it's hard to pin down the belief error. My approach is similar to [AS16] and [Cal16] in that I also classify students into different sophistication levels. However, I reply on additional data about login behaviors from WCPSS application website to help me better identify student types.

2.3 Assignment Mechanism and Data in WCPSS

2.3.1 Assignment Mechanism in WCPSS

WCPSS uses a school choice mechanism to assign K-12 students to magnet schools. Each student is first assigned a base school, which is a non-choice-based assignment determined by their address. Then students may choose to apply to magnet schools. Magnet schools provide special programs to attract students to support the base population of the school. Roughly half of the seats in a magnet school are for base students and the remaining seats are for magnet students who are assigned via the magnet school application process. Students who choose to apply to magnet schools submit their rank ordered list of schools during the application period, which usually lasts two weeks. In 2014, the application website shows the number of students who have ranked each school first and this information is updated continuously during the application period. Students may log into the website at any time during the application window to receive updated information on the number of students ranking each school first. They may change their rankings as many times as they want within the application window. Only their last submitted rankings matter for the assignment. If a student is assigned to a magnet school, she loses her base school and her magnet assignment becomes her assigned school. If a student is not assigned to any magnet school, her base school remains her assigned school. The magnet application process emphasizes to students not to list a school on their application unless they prefer it to their base school.

WCPSS used BM for several years before switching to DA starting in the year 2015-2016. My data are from the academic year 2014-2015, by which time WCPSS had used BM for several years. There's a relatively complicated priority structure in the application process in WCPSS. The priority structure for secondary school applicants in the academic year 2014-2015 is as the following:

- Priority 1: Incoming 6th, and 9th students, who are siblings of current magnet students, wanting assignment to the same school and listing this school first.

- Priority 2: Current magnet students who are rising into 6th or 9th grade that have attended a magnet elementary school or middle school and whose first choice is the magnet middle or high school for their magnet program pathway

- Priority 3: Current magnet students who are rising into 6th or 9th grade and whose first choice is a magnet middle or high school other than their program pathway

- Priority 4: Current base (non-magnet) students in a magnet elementary school or middle school who rising into 6th or 9th and whose first choice is the magnet middle or high school for their magnet program pathway

- Priority 5: Students residing in an area designated as "high-performing" - area whose average test score is among the top one third

- Priority 6: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 120% [6]

- Priority 7: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 115 and less than 120%

- Priority 8: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 110 and less than 115%

[6] A node is WCPSS's terminology for a geographic area that is comparable to a large neighborhood.

- Priority 9: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 105 and less than 110%
- Priority 10: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 100 and less than 105%
- Priority 11: Non-entry-grade students, who are siblings of current magnet students, wanting assignment to the same school and listing this school first.

For elementary school applicants, most of priorities are the same as secondary schools. But there are no pathway priorities for elementary schools applicants and there is a priority that’s specific to the year 2014-2015 (Priority 2 as below), due to some administrative reason. The priority structure for elementary school applicants is as below:

- Priority 1: Incoming kindergarten students, who are siblings of current magnet students, wanting assignment to the same school and listing this school first.
- Priority 2: Current magnet students at school 920532 whose first choice is 920460.
- Priority 3: Students residing in an area designated as "high-performing" - area whose average test score is among the top one third
- Priority 4: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 120%
- Priority 5: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 115 and less than 120%
- Priority 6: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 110 and less than 115%
- Priority 7: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 105 and less than 110%
- Priority 8: Students whose “next” (base) school facility utilization is projected to be greater than or equal to 100 and less than 105%
- Priority 9: Non-entry grade students, who are siblings of current magnet students, wanting assignment to the same school and listing this school first.

Ties among students with the same number of priority points are broken randomly using a lottery number that is not shown to students. In WCPSS, 90% of magnet seats are assigned via BM, while

10% of magnet seats are assigned through a pure lottery that is independent of a student's priority points.[7]

In the academic year 2014-2015, WCPSS had 36 magnet schools: 23 elementary schools, 9 middle schools, and 4 high schools. However, magnet applicants were only allowed to rank up to three schools on their submitted preferences. Further, sibling priority and pathway priorities only applied to a student's first choice. Students' second and third choices were considered after seats were assigned to students' first choices in priority order. But after the first round of the BM algorithm, sibling and pathway priorities are no longer considered. This gives students a stronger incentive to rank first the school where they have sibling or pathway priorities. The next set of priority levels (high-performing node and overcrowded base) still apply to the second or third ranked schools.

For the majority of magnet schools in WCPSS, there is a single program and the assignment process follows BM strictly. For two middle schools, there are multiple programs, which are a form of reserved seats. Specifically, the total capacity of seats at these two schools is divided into a set of seats for academically gifted students and a set of seats for all students. Gifted students may apply to their reserved seats, the unreserved seats, or may list both types of seats on their application (where listing both types of seats would count as two of the three slots on the constrained list). Non-gifted students may only apply to the unreserved seats. Because the sum of the reserved capacity and unreserved capacity equals the total capacity of the school, the WCPSS implementation of BM for these two schools fits within the class of mechanisms considered by [AS16], what they call Report-Specific Priority and Cutoff Mechanisms. As shown by [AS16], all mechanisms within this class have theoretical properties that are conducive to structural estimation. I will explain my structural approach after providing an introduction to my data.

2.3.2 Data and Evidence of Strategic Behavior

My data from WCPSS are richer than what is available in the related literature. I have the following student-level information: base school, demographic information, geographic information, and other information such as educational outcomes (gifted status), English-language learner status, and special education status. Additionally, I have the following information for magnet applicants: preferences submitted to the application website, time of every login to the application website, tie-breaking lottery number, and assignment outcome. I observe applicants' priorities, but I also have information on all components of the priority structure. This allows me to reconstruct priorities, thus I can construct priorities for non-applicants and can construct counterfactual priorities for different implementations of the mechanism (e.g., different mechanisms or different rules about

[7] WCPSS introduced the 10% lottery to encourage more students to participate in the magnet application process. WCPSS policy dictates that non-priority students will not be assigned to a magnet seat (non-priority students are those who do not receive points from any of the enumerated priority levels). As a result, many neighborhoods (i.e., those in a non-high-performing node and whose base is not crowded) were excluded from the magnet process before the introduction of the 10% lottery.

when priorities apply). It is very useful to have data on applicants' lottery numbers, but this is not observed in many other empirical papers. Finally, I observe school-level information on capacity, crowding, and multiple measures of school quality.

In the academic year 2014-2015, there were 155,184 K-12 students in WCPSS. Of these, 5,558 students submitted application for magnet schools in WCPSS. I focus on a sample of 4,999 applicants, which excludes students whose assignment outcomes were nonstandard and observations with data limitations.[8][9] Table 2.1 provides summary statistics for some key demographic variables used in this chapter and compares all students to magnet applicants. The analysis throughout the chapter uses the core sample of 4,999 applicants, with the exception of the calculation of assignment probabilities. With assignment probabilities, the sample includes 4,677 applicants, which excludes 322 students who were seated in the 10% lottery.

The application period was open from January 28, 2014 to February 11, 2014. I use data on visits to the application website to proxy for students' level of sophistication, based on the analysis in [Dur17a]. Following the terminology of [PS08], I classify students as sophisticated if they respond to their assignment probabilities when submitting preferences and as sincere if they do not. In 2014, the WCPSS application website provided information to applicants on the number of students who currently have each school ranked first on their current preference list. Only the final preferences matter for a student's assignment and she can change her preferences as many times as she wishes during the 15 day application period. But, if the preferences submitted thus far are informative for the final preferences, a student can visit the website repeatedly to see the change in relative demand for each school. [Dur17a] show that a student's number of logins reflects her level of sophistication. They show that multiple-login students are 15.1% more likely to receive a magnet assignment, relative to single-login students. [Dur17b] and [Dur17a] take this as evidence that multiple-login students are responding to their assignment probabilities, while single-login students are not.

I follow [Dur17a]'s approach for using students' logins toward an understanding of their level of strategic sophistication. If a student is sincere and thus not responding to her probabilities of assignment, then she does not need to visit the application website more than once. However, if she visits only once, but in the final few days of the application period, then the information shown

[8] I exclude students for the following reasons: students who withdrew their applications (5 students), students admitted to an early college (112 students), students whose assignment was the result of an administrative assignment (304 students), students with missing information on their lottery number (130 students), and students with missing information on their address (8 students). The early college application process runs concurrently with the magnet application process and students admitted to an early college have their magnet application excluded before the magnet assignment is run. Administrative assignments are done for various reasons, such as special educational needs. Of the 4,999 applicants in the remaining sample, there are 2,026 elementary school applicants, 1,748 middle school applicants, and 1,225 high school applicants.

[9] I also have information that allows me to calculate the driving distance of every student to every magnet school. I use ArcGIS, Google Maps, and Mapquest to calculate home-school distances. I drop eight students from the analysis because of address data limitations: five students had addresses that could not be matched to a coordinate and three students had confidential addresses.

on the website about schools' relative demands is a good predictor of final demand. As a result, I classify students as sincere if they visited once and that single visit was outside of the final few days of the application period [Dur17a]. I provide more detail on my use of the login data when I explain the model for estimating preference with heterogeneous sophistication levels.

2.3.3 Ranking and Assignment Statistics

Although students were allow to rank up to three schools, most students did not use all three slots. Among the 4,999 magnet school applicants, 56.4% listed one school, 19.6% listed two schools, and 24.0% listed three schools. Table 2.2 provides the percentage of students at each level (elementary, middle, or high school) by list length. The observation that students do not use all available slots has been previously found (e.g., [Abd09] and [Cal16]). Several explanations have been offered for this behavior, including information costs where students choose not to become informed about all available schools. Note that 73.9% of high school applicants listed one school, even though there were only four available schools. I will present evidence that another explanation also plays a role, specifically that students understood the mechanism and schools' demand well enough to understand that their assignment probability would have been very low at a school had they ranked it second. Further, Table 2.3 shows percentage of students with each of the top four priorities (in the priority structure of secondary schools) in their top ranked school. It is clear that many students top ranked the school where they had a very high assignment probability. Finally, in 2014, 52.6% of applicants were assigned to their first choice through the 90% seats; conditional on receiving an assignment, 90.4% of applicants were assigned to their first choice (which is consistent with the construction of BM).

2.3.4 Assignment Probabilities

As mentioned in Chapter 1, there are two common methods to calculate the assignment probability. Due to the relative large number of students and complicated priority structure, the cutoff approach would be a better choice here, meaning I first need to calculate cutoffs. A cutoff is the number of priority points (including the lottery number) that divides the set of assigned applicants from the set of unassigned applicants. That is, the cutoff at a school is the lowest priority (including the lottery number) a student must have to be assigned to the school. The sample for calculating cutoffs includes 4,677 applicants, which excludes 322 students who were seated in the 10% lottery. The cutoffs are generated for each grade in each school, because each grade has a different capacity and the assignment is run independently by grade. The grade-entry cutoffs and average assignment rates at different ranking positions for each elementary, middle, and high school are shown in Tables 2.4 to 2.6.

In the assignment, WCPSS places a numerical value on each priority level. Priorities for secondary schools are (from highest to lowest) for grade-entry siblings, magnet pathway, magnet non-pathway, non-magnet pathway, high-performing node, overcrowded base, and non-grade-entry siblings. For elementary schools, priorities are (from highest to lowest) for grade-entry siblings, high-performing node, overcrowded base, and non-grade-entry siblings. In 2014, WCPSS randomly assigned a lottery number to each student from the integers from 1 to the total number of applicants. Each student had the same lottery number for all schools. In the counterfactuals, I will refer to this as single tie breaking. I transform the lottery numbers to be uniformly distributed on $(0,1)$.

Beyond the general priority levels described above, WCPSS separates the overcrowded base priority into five separate priority levels in descending order of crowding (from greater than 120% of capacity to between 100% and 105% of capacity). Finally, for elementary schools, in 2014, there was an ad hoc priority level added for one sending school-receiving school pair to handle transfers in response to a magnet program change. Combining these with the above priority description, there are eleven priority levels for secondary schools and nine priority levels for elementary schools. Following [AS16], eligibility scores are required to calculate cutoffs. The eligibility score is constructed as below:

Remark *Eligibility score under BM for WCPSS*

The eligibility score for the rth ranked school s of student i is

$(3-r)*10,000+\sum_{p=1}^{11} 2^{11-p}\left\{\mathbf{P}_{isp}=1\right\}+L_i$, for secondary school students, and

$(3-r)*10,000+\sum_{p=1}^{11} 2^{9-p}\left\{\mathbf{P}_{isp}=1\right\}+L_i$, for elementary school students

where the $\mathbf{P}_{isp}$ is the indicator of whether student i has Priority p in school s, as in the priority structure mentioned above.

Secondary school applicants could receive 1,024 (2^{10}) points for the highest priority (grade-entry sibling), 512 (2^9) points for the second highest priority, and so on. Secondary school applicants could receive 1 (2^0) point for the lowest priority (non-grade-entry sibling). For elementary school applicants, the points range from 256 (2^8) points for the highest priority (grade-entry sibling) to 1 (2^0) point for the lowest priority (non-grade-entry sibling). L_i is the tie-breaking lottery number for student i. I assign 20,000 points for the top ranked school, 10,000 points for the second ranked school, and 0 for the third ranked school, and add this to the sum of the points received from priorities. Once the lottery number is added, this is the eligibility score. Under BM, the cutoff of a school s is the lowest eligibility score among all of those who are assigned to school s.

Given eligibility scores, we can look at the grade-entry cutoff tables. Recall that a cutoff score above 20,000 (10,000) means that a student had to rank the school first (second) to have a chance of being assigned. Cutoffs of 0 refer to underdemanded schools. For elementary schools, there are no underdemanded schools and no schools that were achievable if ranked third. 20 of 23 schools are only achievable if ranked first! Further, for several schools, it is not enough even if students rank them first. For example, the school with the highest cutoff is Douglas E. To be assigned, a student must rank it first and have a very high priority. Specifically, only students with grade-entry sibling priority were guaranteed a seat at Douglas E and even the student at the next priority level[10] were seated only with at least the second lowest crowding priority and a sufficiently high lottery number (i.e., the decimal part of the cutoff). This is consistent with the remaining columns in Table 2.4, where we see that only 30% of students who ranked Douglas E first were assigned to it.

This example helps to demonstrate that both cutoffs and assignment rates tell us the level of competition at each school. In most middle and high schools, it remains true that a student has to rank a school first to be assigned there. But there is more heterogeneity in demand across secondary schools. For middle school, there are 9 schools with 2 schools shown twice, Carnage M and Ligon M, because they have reserved seats that can only be assigned to academically gifted students, denoted as Carnage M AG and Ligon M AG. Five of the other seven schools are only achievable if ranked first, while East Garner M and E Millbrook M were underdemanded. For Carnage M and Ligon M, the unreserved seats are the most competitive seats among all middle schools, while the reserved seats are underdemanded. This makes it clear that WCPSS implemented reserved seats for gifted students in a way that strongly favored gifted students at these two schools by guaranteeing them admission there. The resulting decrease in capacity at Carnage M and Ligon M leads to a lot of competition for the unreserved seats among non-gifted students. Note that Moore Square M also reserves seats for gifted students but does not have a separate program in the application system and instead modifies the priority structure.[11]

For high school, Enloe H is the most competitive high school according to cutoff but Millbrook H is the most competitive school according to assignment rate. This reflects the fact that there are more students who hold the highest priority levels at Enloe H, so tie breaking happens at a higher priority level at Enloe H. In contrast, more students hold the intermediate priority levels at Millbrook H, so the overall assignment rate is lower. Garner H is achievable if ranked second and Southeast Raleigh H is underdemanded. In total, we see a highly competitive environment for elementary school, with less competition in middle school and even less in high school. In high school however, it is still difficult to be assigned to one of the two most popular schools without a high priority and ranking it first. The main message from these results is that students need to be highly strategic if their favorite school is highly demanded and the majority of schools are highly demanded. The result is that the cutoffs are very high for all schools, because the competition at the most popular

[10]The next priority level for Douglas E is high performance note (with 64 priority points), because the special sending school-receiving school pair priority (with 128 points) does not apply to it.

[11]Moore Square M reserved seats for gifted students by guaranteeing assignment for gifted students no matter where they ranked the school. I handle this by giving 30,000 points to gifted students who ranked Moore Square M somewhere on their list. This explains the anomalous result in Table 2.4 that a small number of students who ranked Moore Square M second and third were assigned, even though not all students who ranked it first were assigned.

schools leads some students to rank less popular schools first, they become highly competitive, and so on. This is consistent with the fact that 56.4% of students only listed one school on their submitted preferences. This suggests that sophisticated students realized that, for most schools, they had no chance of assignment if they had ranked it second.

2.4 Model

2.4.1 Utility

There is a set of S schools in the district, denoted as $\{1,2,\ldots,S\}$, that a student can rank and an outside option, denoted as School 0. Students can rank up to K schools, $K \leqslant S$. Student i's utility from being assigned to school s is given by $V(x_{is}, z_s, \epsilon_{is})$, where x_{is} is a vector of observed individual characteristics that may or may not vary by schools, z_s is a vector of school specific characteristics that do not vary by individuals, and ϵ_{is} is the vector of unobserved individual-school characteristics. I include an intercept in x_{is} to capture the school fixed effect, because it measures average utility of the school due to its unobserved characteristics. Any deviation of students' preferences from this average level is captured by the error term, ϵ_{is}, which contains both unobserved individual factors and the heterogeneity in students' preferences over the unobservable school characteristics.

Denote the utility vector of candidate schools as u_i, $u_i = (u_{i1},\ldots,u_{iS})$. The utility of the outside option is normalized to 0, $u_{i0} = 0$. The utility of student i being assigned to school s is

$$u_{is} = V(x_{is}, z_s, \epsilon_{is}) \tag{2.1}$$

where ϵ_{is} is the error term that contains all unobserved variables specific to student i and school s. The vector x_{is} includes characteristics that are constant across schools, such as gender, age, and race. It also includes characteristics that vary across schools, such as the distance of student i's house to school s. The estimates of x_{is} vary across schools. This approach allows, for example, female students to prefer school s to school $s'$and Asian students to prefer school s'' to s. The estimates of schools characteristics z_s are constant across schools, because the estimates denote how much students prefer a particular school characteristic on average, independent of the school itself. Average utilities from unobserved school-specific characteristics are captured by the intercept of x_{is}, which are school fixed effects as explained earlier.

The model makes the following assumptions:

- The error terms of student i at all schools are denoted as a vector ε_i, $\varepsilon_i = (\epsilon_{i1},\ldots,\epsilon_{iS})$. ε_i is i.i.d. across students with a multivariate normal distribution N(0, Σ).
- ε_i is independent from x_{is} and z_s.

- x_{is} and z_s are i.i.d from their population distributions.

- The utility of student i being assigned to school s only depends on her own individual characteristics and school characteristics, and does not depend on others' assignments. This rules out possible network externalities, social interactions, and peer effects.

- Students know their own cardinal preferences over schools and the underlying probability distribution of the preferences for others but not others' realized preferences.

- Student i ranks school s only if $u_{is} \geqslant u_{i0}$, independently of her level of strategic sophistication (sophisticated or sincere).

- Students do not apply to any school, i.e. submit (0, ...,0), if none of the candidate schools has a higher utility than their outside option.

- Sophisticated students' behavior follows a type-symmetric Bayesian Nash Equilibrium, as defined next.

2.4.2 Bayesian Nash Equilibrium

DEFINITION ***Type-symmetric Bayesian Nash Equilibrium***

A Bayesian strategy σ is a *type-symmetric Bayesian Nash Equilibrium*, if every agent of the same priority type and preference order follows the same Bayesian Equilibrium Strategy.

The type-symmetric Bayesian Nash Equilibrium is an extension of the symmetric Bayesian Nash Equilibrium in [Abd11]. Both equilibrium concepts are ex ante, where students face uncertainty from not knowing preferences and reports of other students and uncertainty from the mechanism (unknown tie-breaking lotteries in this case). The difference is that the type-symmetric Bayesian Nash Equilibrium includes priority types in addition to the cardinal preference types in the definition. There are also ex post symmetric Nash Equilibrium concepts in the literature, such as [ES06] and [PS08]. Note that [PS08] prove the existence of Nash equilibria with heterogeneous levels of sophistication and the set of Nash equilibria is equivalent to the set of stable matchings, by using an augmented priority structure.

2.4.3 Heterogeneous Sophistication and Strategies

DEFINITION ***Truth-telling Strategy***

With a truth-telling strategy, students report a ranking that is the same as their true preferences. Students who necessarily follow a truth-telling strategy are called ***sincere***. They submit a ranking $R_i^{\text{truth}} = (r_1, \ldots, r_K)$, where

$$r_1 = \arg\max(u_i, 0), \text{ and} \tag{2.2}$$

$$r_k = \arg\max(u_i \setminus \{r_1, \ldots, r_{k-1}\}, 0), \; \forall k \in \{2, \ldots, K\}.$$

Sincere students do not respond to their assignment probabilities and instead always report truthfully.

DEFINITION ***Expected-utility-maximizing Strategy***

With an expected-utility-maximizing strategy, students submit a ranking that maximizes their expected utility. I refer to these students as ***sophisticated***. They submit a ranking $R_i^{\text{MaxEU}} = (r_1, \ldots, r_K)$, such that $\forall R = (r_1^{'}, \ldots, r_K^{'})$ and $R \neq R_i^{\text{MaxEU}}$,

$$\sum_{k=1}^{K} u_i^{r_k} * \pi_k^{R_i^{\text{MaxEU}}} \geqslant \sum_{k=1}^{K} u_i^{r_k^{'}} * \pi_k^{R}, \text{ and} \tag{2.3}$$

$$r_k > 0 \text{ if and only if } u_i^{r_k} > 0,$$

where $\pi_k^{R_i^{\text{MaxEU}}}$ is the assignment probabilities of student i for school r_k by submitting maximizing expected utility ranking R_i^{MaxEU}, and π_k^R is the assignment probability of student i being assigned to school $r_k^{'}$ by submitting the ranking R.

There might be multiple best responses that follow this strategy, especially when students have a zero assignment possibility for some schools on their list. This usually happens in two circumstances: (1) a student could be assigned with certainty to a school that is on the submitted list and it is not the last school on the ranked list, and (2) at a given position on the ranked list, a student could never be assigned to any of the remaining schools on the ranked list. Both circumstances are common in many school choice settings, including ours. With the estimation approach in this chapter, it does not matter which best response the student follows, as the utility range of zero possibility schools will not be updated during the estimation, as will be explained in more detail. Therefore, I do not need to know the equilibrium selection rule when there is a portfolio of equilibrium strategies. I only need to assume that students follow the same strategy, pure or mixed, as defined in type-symmetric Bayesian Nash Equilibrium.

2.5 Demand Estimation

2.5.1 Estimation of Assignment Probabilities

To estimate students' preference for each school, I first need to consistently estimate assignment probabilities. I need the assignment probabilities of all feasible rankings for each student, because they are needed to find the submitted preferences that maximize a student's expected utility. Given cutoffs, I can calculate the assignment probabilities of each student in each school for each possible ranking. I assume students have the correct beliefs about the distribution of rankings submitted by other students; recall that students not only know how many students have ranked each school first in the current year, but there is also historical information about the distribution of rankings in the previous years. This suggests that sophisticated students could form relatively accurate beliefs of the cutoffs for all schools. Further, students in WCPSS do not know their tie-breaking lottery number. Therefore, I calculate assignment probabilities using cutoffs in the same way as [AS16] and [Kap17], which follows the approach outlined earlier and is done before the realization of the lottery. Note that in WCPSS, only 90% seats were assigned by BM and the remaining 10% were assigned to students' first choice according to their lottery number. As a result, the assignment probability of the first ranked school is adjusted by 10%.

2.5.2 Estimation of Preferences: Homogeneous Sophistication

I estimate preferences by solving a discrete choice model where students maximize their expected utility. The model is non-parametrically identified; see [AS16] for a proof. The identification follows the special regressor approach as in [Lew00], [BH10], and [BH14]. The utility of student i being assigned to school s as in equation 2.1 can be re-written as a form with a special regressor:

$$u_{is} = V(x_{is}^1, z_s, \epsilon_{is}) - x_{is}^2 \qquad (2.4)$$

where $x_{is}^2 \perp \epsilon_{is}$. x_{is}^2 is a special regressor in x_{is} that is linearly separable from the original utility form and depends on both student and school. The coefficient of x_{is}^2 is normalized to 1 or -1 [MR94]. It is set to -1 here, because x_{is}^2 is the home-school distance in the school choice context, which implies that farther distance is associated with disutility. I assume independence of distance to school and unobserved school characteristics, as is done in [AS16] and [Abd15]. However, in those papers, it is likely that some unobserved school characteristics (such as school quality) determine school preferences and residential choice simultaneously, contrary to the independence assumption. Here, the magnet school environment of WCPSS is more consistent with the independence of distance and unobserved school characteristics. I argue that students are unlikely to make their residential choice as a function of distance to a given magnet school, because residential area determines one's base school and does not directly affect one's magnet assignment probability. Recall that around 50% of the seats at each magnet school are reserved for base students who live in the school's neighborhood. Thus, if a student makes her residence choice to increase her assignment probability at a given magnet school, she can simply move into the school's neighborhood and be assigned to the school as a base student. Such a student would not appear in my data, because she would not be a magnet applicant, having been assigned to her preferred school via the base assignment. Finally, the normal distribution of the error term also satisfies the identification requirement in [AS16].

Although the model is non-parametrically identified, it is common to use a parametric model for point estimation in finite samples. The parametric utility of student i being assigned to school s is:

$$u_{is} = \beta_s x_{is} + \delta z_s - d_{is} + \epsilon_{is} \tag{2.5}$$

$$u_{i0} = 0$$

where x_{is}, z_s, and ϵ_{is} are as described in Section 3 and d_{is} is the distance from school s to the house of student i. The estimates of x_{is} vary by schools and the estimates of z_s do not depend on school. Suppose there are J different characteristics in x_{is} and Q different characteristics in z_s, then equation 2.5 can be written as:

$$u_{is} = \sum_{j=1}^{J} \beta_{sj} x_{isj} + \sum_{q=1}^{Q} \delta_q z_{sq} - d_{is} + \epsilon_{is} \tag{2.6}$$

The baseline model is with homogeneous sophistication, assuming all the students are fully sophisticated. Although this is not a completely realistic assumption, it is a natural starting point for estimation. There are very few students (60 out of 2,026 elementary school applicants, no middle or high school applicants) whose rankings cannot be rationalized.[12] I exclude these students from the estimation of the homogeneous sophistication model. Note that 98.8% of rankings can be rationalized in these data, which is higher than in [AS16] (92%).

2.5.2.1 Characteristics With School Varying Estimates Only

First, I follow the setting of [AS16] and only include the vector x_{is} whose estimates vary across schools. It includes student-level characteristics that may or may not vary across schools. It also includes the school-specific dummy (intercept). I refer to this as the school fixed effect because it measures the average utility of the school due to its unobserved characteristics.

$$u_{is} = \sum_{j=1}^{J} \beta_{sj} x_{isj} - d_{is} + \epsilon_{is} \tag{2.7}$$

$$u_{i0} = 0$$

Because this setup is the same as the model in [AS16], I use the same estimation setup:

[12] A ranking that cannot be rationalized involves a dominated action such as the following: ranking a school third that has zero assignment probability when ranked third but positive assignment probability when ranked second, while the school ranked second has a zero assignment probability when ranked second.

$$u_i = X_i\beta - D_i + \varepsilon_i \tag{2.8}$$

X_i is a $S \times SJ$ block-diagonal matrix that is constructed by placing the J-row vector covariates $x_{is} = [x_{isj}]_{j=1}^{J}$ in each of the J blocks, as shown below:

$$X_i = \begin{bmatrix} x_{i1} & 0 & \cdots & 0 \\ 0 & x_{i2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & x_{iS} \end{bmatrix}$$
$$= \begin{bmatrix} [x_{i11}\, x_{i12} \cdots x_{i1J}] & 0 & \cdots & 0 \\ 0 & [x_{i21}\, x_{i22} \cdots x_{i2J}] & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & [x_{iS1}\, x_{iS2} \cdots x_{iSJ}] \end{bmatrix}$$

where S is the total number of schools and J is the total number of student-level characteristics.

β is a $SJ \times 1$ vector:

$$\beta = \left[\beta_1 \ldots \beta_S\right]^T = \left[\beta_{11} \ldots \beta_{1J}\; \beta_{21} \ldots \beta_{2J} \ldots \beta_{S1} \ldots \beta_{sJ}\right]^T$$

D_i is a $S \times 1$ vector of distances from the house of student i to all the S schools and $\varepsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{iS})$. The Bayesian estimation uses Gibbs Sampler with data augmentation following [MR94]. Utilities are updated following [AS16]. In particular, the initial values of utilities for each student need to satisfy the maximization of expected utility, given the assignment probabilities for all possible rankings. Then the utility for each school follows a truncated normal distribution, whose upper bound and lower bound are updated in the utility maximization process by solving linear inequalities. I provide a detailed estimation process in the appendix, which is based on [AS16] but updates their process in several ways as explained in the appendix.

The estimation results for elementary, middle, and high school applicants are shown in Tables 2.7, 2.8, and 2.9.[13] When I estimate preferences of middle schools, the two gifted programs are considered together with regular programs for the same school. Once students are admitted, there is no separation of students seated via the gifted program versus other students. Therefore, there should be no difference in their preferences of the school. However, the assignment probabilities

[13]The estimation results are robust in the following dimensions: alternative initial values for utilities, alternative prior distributions, and alternative numbers of loop and burn-ins. I find that 15,000 loops and 5,000 burn-in are sufficient for high school estimation. For middle and elementary school estimation, longer loops are desired. The results shown here use 100,000 loops with 50,000 burn-in.

will differ across programs.

2.5.2.2 School Level Characteristics With School Constant Estimates

[AS16] only include characteristics whose estimates vary across schools, as shown in the preceding model. Besides the student-level characteristics, the model also measures the school fixed effect, which is the average utility of the school due to its unobserved characteristics. However, school districts are often interested in how certain school-level characteristics may determine students' preferences. I expand the model to include school-level characteristics whose estimates do not vary across schools, as in the model in equation 2.5. The estimation setup is:

$$u_i = X_i\beta + z\delta - D_i + \varepsilon_i \tag{2.9}$$

where z is a $S \times Q$ matrix whose sth row is Q school-level characteristics of school s and δ is a $Q \times 1$ vector. During estimation, I stack X_i vertically for each student i to form a $SN \times SJ$ matrix called X, where N is number of students. I also stack z vertically for each student to form a $SN \times Q$ matrix, even though z is the same for all students, because I jointly estimate $[X, Z]$ in order to update β and δ in one step. See the appendix for more details.

The school-level characteristics used for elementary schools and middle schools are a dummy variable for traditional calendar (equals 0 if the school uses a year round calendar), teacher turnover rate, and School Performance Grade (SPG) score.[14] The school-level characteristics used for high schools are teacher turnover rate and SPG score; traditional calendar is omitted because most high schools use a traditional calendar. SPG score measures the general academic performance and it is widely used as a main index for school quality. Estimation results for elementary, middle, and high school applicants are shown in Tables 2.10, 2.11, and 2.12. Traditional school calendar positively affects utility on average for elementary schools but negatively for middle school. The standard deviation of traditional calendar for both levels is very large, which means there is preference heterogeneity with some students strongly preferring a year round calendar. The estimates of the rate of teacher turnover are negative for middle school and high school. Although it is positive for elementary school, it is not significantly different from 0. Estimates of SPG at all three levels are positive, which means most parents prefer a school with good academic performance. Note that the coefficients of school-level characteristics converge much faster than the coefficients of student-level characteristics, which suggest that school-level characteristics are well identified.

This model allows me to explain how much of the variation in average utility across schools is explained by school-level characteristics. The utility associated with each school is modeled as having a scalar component (the intercept) and additional components for student-level characteristics

[14] For the remainder of the chapter, I use the term "school-level characteristics" to refer to the characteristics in vector z whose estimates do not vary across schools, excluding the school fixed effect.

(e.g., race) and school-level characteristics (e.g., test scores). Without school-level characteristics, the school fixed effects capture most of the variation in utilities across schools. Schools can be ranked simply by comparing the estimates of their fixed effects when there are no school-level characteristics. After adding school-level characteristics into the model, the school fixed effects explain much less of the variation in utility across schools and observed school-level characteristics play an important role. Total school-level characteristics (including observed characteristics in z and unobserved characteristics measured by intercept) explain more than 90% of the variation in utility, excluding the utility associated with distance from the student to the school. This implies that student-level characteristics explain a small fraction of the variation in average utility across schools.

In addition to estimating differences in the average utility across schools, the model also provides estimates of the differences in utility of a given school across students. This goes beyond explaining which schools are highly or lowly demanded, on average, to explain which schools are preferred by students with certain characteristics. Student-level characteristics include observed characteristics (e.g., race, gender, and status as being academically gifted) and unobserved characteristics. The model attributes differences in utility of a given school across students due to unobserved characteristics to the standard deviation of the utility. This allows heterogeneity in students' preferences due to unobservable factors. The results match anecdotal evidence concerning preferences. For example, Southeast Raleigh High is estimated to have a very large variation in utility across students. I find that this school is popular among students who want technical training classes, but not popular among students who want college-prep classes. This is consistent with the finding that academically gifted students have lower utilities for this school than students who are not academically gifted. This taste heterogeneity is estimated as a variation around the average utility level.

2.5.2.3 Random Coefficient Model

Beyond including school-level characteristics into the model, there are additional sources of heterogeneity in tastes over school-level characteristics. The estimation results show that the standard deviations for all school-level characteristics are very large. This intuitively reflects unobserved heterogeneity in students' preference over school-level characteristics. This can be accommodated with a random coefficient model, where preferences follows a distribution and different students locate at different points of the distribution. If the random utility coefficients model is used with cross sectional data, the estimates are the parameters of the preference distribution. To get student-level estimates of preferences, I need panel data. The random coefficient model can be estimated with panel data using a Bayesian hierarchical model as shown in [MR94]. [Ros96] and [AR99] extend this approach to get individual-level estimates in a model with heterogeneous tastes. All these papers use panel data. However, it is difficult to get panel data for school assignment, because we cannot observe repeated decision of students over same schools. It is computationally difficult to estimate a random coefficient model with cross sectional data using the Bayesian approach. If we assume the estimates $\beta_i \sim N(b,\ W)$, then posterior distribution of β_i would be

$$K(b,\ W|U) \propto \prod_{i=1}^{N} L(R_i|b,\ W)k(b,\ W) \tag{2.10}$$

where $U = (u_1, \ldots,\ u_N)$; $L(R_i|b,\ W)$ is the probability of student i submitting ranking R_i given b and W ; and $k(b,\ W)$ is the prior on b and W (normal for b times inverse Wishart for W). In the process, we have to update b and W. Differently from with panel data, β_i is not identified with cross sectional data and it cannot help update b and W in the EM algorithm. With cross sectional data, we have to draw directly from $K(b,\ W|U)$ with the MH algorithm, which is extremely slow [Tra09]. For each iteration of the MH algorithm, it is necessary to calculate the right-hand side of equation 2.10, which requires a simulation of $L(R_i|b,\ W)$ (which has no closed form and can only be estimated by simulation) for each student. The computational difficulties prevent me from applying the Bayesian approach to the random coefficient model with cross sectional data.

2.5.3 Estimation of Preferences: Heterogeneous Sophistication

In the above model, I assume that students are fully sophisticated and submit the ranking with highest expected utility. As mentioned earlier, around 1.2% of students submitted rankings that cannot be rationalized. [Cal16] do not impose the type of these agents. They estimate a mixture model where the identification does not rely on the rankings with obvious mistakes. [AS16] build a mixture model with heterogeneous types of students where the types follow a binomial distribution. They start with a uniform prior (Beta (1,1), 50% probability of being sincere and 50% probability of being sophisticated) for the student type, initially set every student as sincere, then update the percentage of sincere students via Markov Chain Monte Carlo. Students whose types are indeterminate are randomly assigned a type from the binomial distribution. The randomness introduced by this approach might be problematic and indeed the percentage of sincere students estimated by [AS16] is high relative to the work in the related literature. Further, in my data, this approach is quite sensitive to the prior distribution for types.

My innovation is to use novel data to aid in the identification of student types. The most closely related approach in the literature is that of [Kap17], who conduct a survey to estimate beliefs regarding assignment probabilities. I use data from the application website where students submitted their preferences, based on the analysis in [Dur17a]. In 2014, the WCPSS application website provided information to applicants on the number of students who currently have each school ranked first on their current preference list. Only the final preferences matter for a student's assignment and she can change her preferences as many times as she wishes during the 15 day application period. But, if the preferences submitted thus far are informative for the final preferences, a student can visit the website repeatedly to see the change in relative demand for each school. Using these auxiliary data, my estimation of student types has better properties with respect to empirical identification.

I apply a similar Gibbs Sampler process as in the homogeneous sophistication model, given the type of each student (sophisticated or sincere). I define the type of student depending on the number and timing of their logins to the application website. I define a student as sophisticated if this student logged in the website more than one time or the student's last login happened in the last two days of the application window. I argue if a student is sincere, the timing of submission and frequency of logins to the application website should not matter. In contrast, sophisticated students should login several times to check the updates of how many students have ranked their preferred schools first. If a student logged in the website once but within the last two days, she is treated as sophisticated, because the information shown on the website about schools' relative demand is fairly precise at that point. Finally, if a student logged in the website once but has a "dominating" priority at the first ranked school, she is also treated as sophisticated, because these students are guaranteed an assignment at the school they ranked first and might only need to log in once because that school's relative demand does not matter given their high priority. An example is a grade-entry student with sibling priority applying to their sibling's school. She knows that she will be admitted to the school if she ranks it first, so logins are less relevant for her. Following this definition, I find the percentages of sincere students among elementary, middle, and high school applicants are 16.1%, 13.0%, and 14.8%, respectively.

I estimate the heterogeneous sophistication model on middle and high school applicants and the results are shown in Tables 2.13 and 2.14. The demand estimate for high school are quite similar for the heterogeneous sophistication model and the homogeneous sophistication model. For example, the relative popularity of schools is unchanged across the two models. This is likely because 73.9% high school applicants in my sample only submitted one school and the estimates suggest that the average utility of schools are similar for sophisticated and sincere students. The changes in the middle schools estimates across the heterogeneous and homogeneous sophistication models are meaningful, especially the estimates of gifted students, non-grade-entry students, and unobserved school fixed effects. Using anecdotal evidence regarding schools' popularity, I argue that the estimates of the heterogeneous sophistication model are more intuitive. For example, gifted students have significantly positive utility for Carnage Middle and Ligon Middle under the heterogeneous sophistication model, while the effect is negative under the homogeneous sophistication model. Because these two schools focus more on academic achievement and are more academic-gifted students concentrated, I expect gifted students to prefer these two schools, ceritas paribus.

2.6 Ranking Behaviors

I estimate each student's cardinal utility for each school and uses these estimates to analyze students' ranking behavior. This is useful because it provides evidence for strategic behavior in the field. I study the ranking behaviors of middle and high school applicants using the estimation results from the heterogeneous sophistication model. For elementary school applicants, I use the estimation results from the homogeneous sophistication model with school characteristics.

I categorize ranking behavior into six strategies: truth-telling, truncated truth-telling, skip-the-top, skip-the-middle, flipping starting in position 1 (flip1), and flipping starting in position 2 (flip2). Truth-telling means a student reports a ranking that is exactly the same as their true preference, up to the constraint of the mechanism (listing up to three school). Truncated truth-telling means a student reports truthfully but reports a truncated version of their true preferences. For example, truncated truth-telling involves a student who truly prefer two schools to her base school but only ranks her favorite school and leaves the remaining two slots empty. Skip-the-top means a student does not report her true favorite school first and instead reports a school somewhere in the middle of her true preferences. The rest of reported schools are in the same order as her true preferences, but certain schools might be skipped. Skip-the-middle means a student reports her true favorite school first but then omits her next truly preferred school and instead reports schools that are lower in her true preferences. Flipping starting in position 1 means a student does not report her true favorite school first and the second school she ranks is in fact more preferred to the first ranked school. Flipping starting in position 2 means a student reports the first school truthfully, but the second ranked school is in fact lower in the true preference list than the third ranked school.

Suppose a student's true preferences are schools $A-B-C-D$. Truth-telling is reporting $A-B-C$, truncated truth-telling is $A-B$ or A, skip-the-top is B, $B-C$, $B-C-D$, $B-D$, C, or $C-D$, skip-the-middle is $A-C$, $A-D$, $A-C-D$, or $A-B-D$, flip1 is $B-A$, $B-A-C(or D)$, $C-A(or B)$, $C-A-B(or D)$, $C-B$, etc., and flip2 is $A-C-B$, $A-D-B$, or $A-D-C$. Finally note that a student who flips in both positions 1 and 2 (such as the $C-B-A$ in the example above) will be categorized as flip1.

The percentage of the six types of ranking behaviors are shown in the Panel A in Table 2.15. The Panel B groups some of above behaviors that follow a similar pattern. The rates of truth-telling and truncated truth-telling are lowest for elementary schools and highest for high schools. As I mentioned previously, the application process for most elementary schools is very competitive, which leads to heavier manipulation. As shown in Panel B in Table 2.15, 93.7% middle school applicants ranked the first school truthfully. This is consistent with my estimation results that there is greater taste heterogeneity for middle schools. Lower correlation in preferences lowers competition and more students could rank their favorite school truthfully. In all three grade level, truncated truth-telling is the most common ranking behavior. I argue that this reflects a high degree of understanding of the mechanism and the competitive environment among students in my sample. The modal observation (and the majority of observations in middle and high school) involve a situation such as the following: a student's true preferences are $A-B$. She reports A, while had she reported $A-B$, her assignment probability at school B when ranked second is zero. I conclude that the student is not using the second slot on the submitted list because she has a good understanding of the assignment probabilities.

2.7 Counterfactual Analysis

The estimation results allow me to study alternative designs of a school choice mechanism. In this section, I consider the following four design elements: type of mechanism, constrained or unconstrained lists, single or multiple tie-breaking lotteries, and reserved seats or not. The reserved seats in WCPSS favor academically gifted students (in the form of a separate program and assigning a higher priority). I focus on two effects of these factors, efficiency and fairness (i.e., elimination of justified envy). I compare the following mechanisms: BM, DA, and TTC. For BM and DA, I have both constrained lists (rank up to three schools) and unconstrained lists. For TTC, I only have the unconstrained version. For all the mechanisms, I compare the difference between a single lottery across all schools (single tie breaking) and a separate lottery for each school (multiple tie breaking).

In WCPSS, 10% of all magnet seats are assigned solely based on students' tie-breaking lottery numbers to students who rank this school first. To make a more straightforward comparison with other school districts, I exclude the 10% lottery from my counterfactual analysis. All seats at each school are included (i.e., the 10% lottery seats are treated as regular seats) and all students are included (i.e., students whose assignment in 2014 was as a result of the 10% lottery are treated the same as all other students). As a result, the constrained BM analysis in this section is a counterfactual implementation of the WCPSS mechanism without the 10% lottery. The inputs to the counterfactual analysis are students' true preferences (estimated in the previous section), students' priorities (observed in the data), and schools' capacities (observed in the data). For middle and high schools, I use the estimation results from the heterogeneous sophistication model with school characteristics. Under manipulable mechanisms, I find each student's submitted preferences as follows: sincere students follow the truth-telling strategy (up to the constraint of the list if applicable) and sophisticated students follow the expected-utility-maximizing strategy. For elementary students, I use the estimation results from the homogeneous sophistication model with school characteristics and assume all students are sophisticated. Further, the counterfactual analysis for elementary school presents results for "unconstrained" BM that is in fact constrained to ranking up to eleven schools. This is done for computational reasons with respect to finding students' expected-utility-maximizing strategies under unconstrained BM.[15] For all counterfactuals, I use the bootstrap

[15]With unconstrained lists under BM, the number of possible rankings students might submit grows too large for

method, re-sampling with replacement.

2.7.1 Efficiency

I first compare mechanisms and designs in terms of efficiency. Matching theory demonstrates that, among strategyproof mechanisms, TTC is efficient, while DA is not. BM cannot Pareto-dominate DA with complete information [ES06], but the comparison is more nuanced with incomplete information. (See [Abd11], [Mir08]), and [Tro12] for detailed analyses.)

To compare efficiency, I use a cardinal measure (the utility of the average student) and two ordinal measures (the percentage of students assigned to the true favorite, Rate1, or one of their three favorite schools, Rate3). The utility of the average student assigns zero utility to all unassigned students, so two mechanisms can differ in their average utility both through assigning different numbers of students and to assigning students to more or less preferred schools. The mechanisms considered are TTC, constrained BM, unconstrained BM, constrained DA, and unconstrained DA. Constrained versions of the mechanisms restrict students to submitting at most three schools. I first discuss the efficiency differences associated with difference design choices for the assignment procedure. Second, I discuss the differences observed for sophisticated and sincere students with different designs. In particular, I want to know whether sophisticated students gain from manipulable mechanisms at the expense of sincere students, which has been an importance concern with mechanisms such as BM [PS08]. The efficiency comparison are shown in Tables 2.16 to 2.19.

Before discussing the results, I discuss the observed assignment results under BM in 2014-2015. The nature of competition for seats varies greatly across elementary, middle, and high school in WCPSS. This provides useful variation that helps me understand how mechanisms and their design compare in different settings. However, it also implies that some of the results differ for different grade levels. There are 23 elementary schools and the observed rate of assignment in elementary school for grade-entry students is 56.5%. As shown by the cutoffs in Table 2.4, the majority of elementary schools are only obtainable if a student ranks it first. There are 9 middle schools and the assignment rate in sixth grade is 67.0%. As shown by the cutoffs in Table 2.5, there are a set of highly competitive schools and a small number of underdemanded schools. Further, middle school is the only grade level with reserved seats, so I analyze reserved seats for middle school only. Finally, there are 4 high schools and the assignment rate in ninth grade is 76.3%. The cutoffs for high school suggest large differences in demand across schools, as with middle schools. An important observation in the high school assignment is that 72.8% of high school applicants only listed one school on their submitted preferences.

computation, which prevents me from finding the equilibrium reports when students can report full lists under BM. Note that this only affects elementary schools and the majority of elementary school students do not have positive utility at more than eleven schools, which implies that the mechanism is completely unconstrained for these students.

2.7.1.1 Alternative Mechanisms and Design Choices

In this subsection, I ignore differences across students with different level of sophistication and discuss how the design of a school choice mechanism affects efficiency. Starting with high school, I do not observe any significant differences in efficiency across different mechanisms or for different tie-breaking rules as indicated in Table 2.19. This is consistent with the high school estimation results, where I find that there is substantial heterogeneity in preferences among high school applicants and a set of students prefer underdemanded schools. There are significant differences across designs in elementary and middle schools. In Table 2.16, TTC is more efficient than all other mechanisms except unconstrained BM. Unconstrained BM has slightly higher average assignment utility and assignment to one of students' top three preferred schools than TTC, but there are significantly more students are assigned to their favorite school under TTC than under unconstrained BM. In Table 2.17, TTC is the most efficient mechanism in terms of both average utility and assignment rates to preferred schools. As far as I know, [Abd17] and [Cal16] are the only papers that have empirically compared the efficiency of TTC to other mechanisms. [Abd17] find TTC is only slightly more efficient than DA in term of assignment rates. The efficiency difference I find is slightly larger, especially the assignment rate to their favorite schools; I argue that this is because there more competition for seats in WCPSS relative to the districts studied by [Abd17]. They found that 18% of students were unassigned in New Orleans and 8% of students were unassigned in Boston. The rates of unassigned students in WCPSS are much higher: for elementary school applicants, 54% of students are unassigned. Because competition for seats is more intense in WCPSS, the efficiency gain from using TTC is slightly larger than in [Abd17]. [Cal16] in fact find larger efficiency difference between TTC and BM or DA than what I have found. They find there's an over 50% gain by switching from DA to TTC and the efficiency gain is even larger for the non-strategic households. Given the limited empirical evidence from TTC, I do not know the size of TTC's efficiency advantage in other settings. But these results suggest that it can be considerable. In middle school, the effect size is the largest: TTC assigns 4 percentage points more students to their favorite relative to unconstrained DA, which is an 11.17% increase; however, there are very small differences in average assignment utility, assignment rate to one of the top three preferred schools, and total assignment rate. As predicted by the theoretical literature, using a mechanism other than TTC can come at an efficiency loss.

In all grade levels, BM is weakly more efficient than DA, but these differences are relatively small. The largest efficiency differences between BM and DA are in middle school, where BM generally assigns 2.5 percentage points more students to their favorite school. However, BM and DA generally assign similar numbers of students to one of their three favorite schools, especially under constrained versions. Further, the difference in cardinal efficiency between BM and DA always suggests an advantage for BM that is small (e.g., the efficiency comparisons in Panel A of 2.17 are all approximately 1% higher efficiency for BM over DA). The theoretical literature has suggested that BM might be more efficient than DA and some empirical papers have found this to be the case (e.g., [Cal16] and [AS16]). Relative to other empirical results, the efficiency differences for BM and DA here are smaller. A conclusion that could be drawn in these results is that efficiency loss is not a reason to avoid switching from BM to DA for school districts who are considering such a switch.

Moving to other comparisons, I am interested in understanding the effects on efficiency from constraining a mechanism to only allow students to report three schools, using single or multiple tie-breaking lotteries, and reserving seats to achieve academic or diversity goals. Overall, constrained versus unconstrained lists has a relative small effect on efficiency and this is also true for single versus multiple tie-breaking lotteries. In contrast, the outcomes of these mechanisms are very different if the school district implements reserved seats or not.

For constrained lists, the unconstrained version of BM is always weakly more efficient than constrained BM. Unconstrained DA is more efficient than constrained DA in elementary and high school, but less efficient in middle school. However, all of these differences are quite small. For tie breaking, the differences in efficiency for a given mechanism with single versus multiple lotteries is very small. There are no meaningful differences at any level for any mechanism. The theoretical matching literature has suggested that multiple tie breaking can have efficiency losses, but I see no evidence of this.

For reserved seats, the efficiency differences are large. Table 2.18 lists results both with and without reserved seats for all middle school students. As we can see from Table 2.18, average utilities with reserved seats are lower than the average utilities without reserved seats under all mechanisms, while the assignment rates to favorite schools are higher with reserved seats. Gifted students are much more likely to be assigned to one of the gifted programs, which is intuitive, and the preference estimation reveals that these students in fact truly prefer these schools. This results in a large increase in the assignment rates. But the increased competition at those three schools causes some non-gifted students who prefer those schools to avoid them. This increases competition at other schools, which lowers average utilities with reserved seats compared to the assignment with no reserved seats.

2.7.1.2 Effects of Student Sophistication Type

As explained earlier, all elementary schools students are treated as sophisticated, so I compare efficiency by student sophistication in middle and high schools only. Tables 2.17 and 2.19 shows that the average utility and assignment rate to favorite schools are both lower for sincere students than the sophisticated students. The difference is larger in high school. Given my approach for classifying a student as sincere or sophisticated, it is difficult to interpret the relative magnitudes of each group's average utility. Instead, I will focus on differences in efficiency for a given type across different mechanisms and different design choices.[16]

I am interested in the heterogeneous effects of the choice of mechanism on sophisticated versus sincere students. [PS08] argue that a strategyproof mechanism will “level the playing field” by reducing the advantage that can be gained by strategically manipulating a manipulable mechanism. To empirically test this, I start from the baseline of constrained BM. This is the mechanism that is most commonly used in practice and economists have frequently recommended that districts switch to a strategyproof mechanism (and to a mechanism with unconstrained lists). Relative to constrained BM, switching to DA (unconstrained or constrained) reduces utility and assignment rates for sophisticated students, while the effect is less clear for sincere students. Switching to DA increases average utility for sincere students and increases their assignment rate at one of their three favorite schools but reduces their assignment rate at their favorite school. This is consistent with other evidence that manipulable mechanisms have assignment outcomes that are more “risky” in the sense of higher assignment rate to the favorite school but also higher rate of unassignment, relative to strategyproof mechanisms (e.g., [Cal16] and [AS16]). The increase in average utility for sincere students in the switch from BM to DA, along with the decrease in the favorite-school assignment rate, implies that some sincere students gain from the manipulation of sophisticated students, but sincere students whose favorite schools are highly competitive schools are likely to be hurt by reporting truthfully.

The conclusion I draw from Tables 2.17 and 2.19 is that switching from BM to DA will in fact level the playing field. If I take the assignment rate at one of the top three favorite schools as a good measure of efficiency, switching from BM to DA slightly reduces assignment rates for sophisticated students and substantially increases assignment rates for sincere students. This is exactly what [PS08] predicted and what economists have argued to school districts when convincing some districts to switch away from BM. That said, the efficiency results for TTC might suggest that the focus on the gains from switching to DA should also include a stronger focus on other strategyproof mechanisms. In middle school, all students gain from switching to TTC. In high school, the comparison is less favorable to TTC, but there is not much evidence that the choice of mechanism has much effect on efficiency. It remains to be seen whether school districts are willing to use TTC, given the results observed in New Orleans and the discussions surrounding the switch to DA in Boston where priority “trading” was viewed negatively by the school board.

Economists have also argued that school districts should allow students to submit lists that are either unconstrained (when possible) or increase the number of slots on the list (when the number

[16] A partial explanation for sincere students’ lower average utility is in the way I defined sincere and sophisticated students, where students who logged in once but had a very high priority at the top ranked school were counted as sophisticated. This implies that sophisticated students are slightly more likely to have a very high priority at the reported first choice. To the extent that the estimation reveals these top ranked schools to be these students true favorite schools, the estimation results in part reflect this feature in the average utility differences between sincere and sophisticated students.

of schools or school programs is very large). Tables 2.17 and 2.19 present conflicting results on the efficiency effects of constrained lists. In high school, constrained lists generally reduce efficiency, but the effect is small. Constraints on BM and on DA have similar effects. In middle school, constrained lists have ambiguous effects depending on mechanism and type of students. The differences are generally small though. I conclude that constrained lists do not have a large efficiency loss. In fact, the only differences I view as moderately large are for sincere students in middle schools. For them, constrained BM increases their assignment outcomes (higher average utility and assignment rates). The interpretation of this result is that the strategic behavior of sophisticated students with constrained lists expands the set of achievable schools for sincere students in middle school. However, the aggregate differences are not large enough to argue for constraining students' lists on efficiency grounds.

2.7.2 Fairness

Fairness is the elimination of justified envy, where justified envy implies that a student is assigned to a school that you prefer to your assignment but has lower priority than you at that school. The empirical school choice literature has extensively discussed efficiency but there has been less focus on fairness. This is likely due to the fact that DA is a fair mechanism, that is, DA with unconstrained lists will necessarily have zero instances of justified envy. To examine the tradeoffs of several mechanisms, including TTC, I examine fairness here. For justified envy, I report the total number of instances of justified envy, the number of students who have justified envy of at least one other student (i.e., "jealous" students), and the number of students who are envied by at least one other student (i.e., "envied" students). Jealous and envied students are also shown as a percentage of all students by grade level. Finally, I report the average number of instances of justified envy per jealous student. I first calculate these results for each bootstrap sample, and then calculate the average across all the bootstrap samples. The bootstrap average results are shown in Table 2.20 to 2.23. In terms of jealous students, the percentage of jealous students can be interpreted as the extensive margin of fairness, while average envy per jealous student can be interpreted as the intensive margin of fairness.

Papers on school choice have primarily measured fairness by the number of instances of justified envy and the number of students who have justified envy of at least one other student, such as [Pat16], [AS16], and [Abd17]. I use the term "jealous students" for the latter concept and additionally include the number of envied students in my comparisons. The rationale for focusing on envied students is as follows: suppose a student who is very low in the priority order at school A is assigned there by TTC because she "traded" her priority at another school to be assigned to A, which was her favorite school. If she is low in the priority order, then many students could have justified envy of her. However, only one of those students with justified envy could be assigned to the seat that she holds.

Thus, counting the number of students who envy her can be considered over-counting the extent to which the mechanism is "unfairly" assigning seats. The rationale for focusing on (what I call) jealous students is more practical. Every student who can deduce that she has justified envy can complain to the school district or file a lawsuit in an extreme case. This latter rationale has played a large role in arguments that TTC is poorly suited to school choice. I include multiple measures of fairness to provide a comprehensive comparison.

Throughout this discussion, I compare the four mechanisms that have some justified envy, which ignores unconstrained DA. As always, if the complete elimination of justified envy is required, none of these mechanisms can improve on unconstrained DA. Among the other four mechanisms, TTC has the highest number of justified envy instances and most jealous students, while it has the fewest envied students. The average justified envy instances per jealous student is generally lowest with TTC except in middle school without reserved seats. In total, TTC assigns a small number of students to schools at which they have low priority, but these assignments result in a lot of jealous students.

The percentage of jealous students in My sample is higher than [Abd17], but they do not provide statistics on the number of envied students. In [Abd17], around 13% of students have justified envy (jealous students) under TTC using data from New Orleans, while around 7% of students have justified envy under a counterfactual implementation of TTC using data from Boston. I conjecture that I find higher rates of jealous students for two reasons: relative to New Orleans and Boston, (1) the magnet school assignment in WCPSS is more competitive and the correlation among preferences is higher and (2) there are more levels in the priority construction in WCPSS, which implies that there are more levels of students along which a student can envy or be envied by.

If minimizing the number of jealous students is the preferred fairness metric, BM is much more fair than TTC but still meaningfully worse than constrained DA.[17] In contrast, if minimizing the number of envied students is preferred, BM is the least fair mechanism. Further, the effect of constrained lists on fairness is ambiguous. In elementary school, I find more jealous students and more envied students with unconstrained BM than constrained BM. This suggests that strategic manipulations reduce justified envy in the highly competitive elementary school assignment. However, in middle and high schools, all unconstrained versions of BM and of DA has less justified envy than constrained versions according to all fairness metrics.

Constrained DA is not a fair mechanism, but there is not a lot of evidence telling us how much justified envy exists when DA is implemented with constrained lists. My results suggest that the fairness cost of constraints with DA depend, again, on the measure of fairness. Constrained DA generally has a small number of jealous students but a moderately large number of envied students. As a result, constrained DA loses a lot of DA's appeal in terms of fairness if the district is worried

[17] Fair is a binary concept in that an assignment is either fair or not. However, I loosely speak of comparisons in terms of frequency of justified envy as tells me the degree of fairness.

about jealous students but not if the district is worried about envied students.

On other design choices, the tie breaking rule (single lottery for all schools or a separate lottery at each school) does not have a large effect on the extent of justified envy. This is consistent with the efficiency results and overall tie breaking does not seem to have large effects. Finally, for reserved seats, justified envy is moderately lower when some seats are reserved. But this is in some sense mechanical because a student cannot have justified envy of a student of a different type. In the case of reserved seats in WCPSS, academically gifted students have reserved seats at three middle schools. They are more likely to be assigned there and non-gifted students cannot envy those seats because non-gifted students cannot hold them. Further, reserving those seats for gifted student lowers competition at other middle schools, which further reduces the amount of justified envy. Finally, the patterns of justified envy across mechanisms do not differ with or without reserved seats, only the level of justified envy.

2.8 Conclusion

In this chapter, I use the Bayesian approach and a multivariate probit model to estimate the demand of schools under BM with the WCPSS magnet school application data. I estimate preference using three different models: two homogeneous sophistication models (with and without school characteristics) and one heterogeneous sophistication model. I identify the types of students using login data and estimate preference given the type in heterogeneous sophistication model. I find that the school level observed and unobserved characteristics explain more of a school's demand than individual level characteristics. I classify students' strategic behavior in submitting their preferences over schools and find a lot of manipulation. The types of manipulations I observe provide suggestive evidence that applicants have a good understanding of assignment probabilities.

I conduct several counterfactual analyses on different designs of a school choice mechanisms to examine the effects on efficiency and fairness regarding to the choice of mechanism, the existence of constrained lists, tie-breaking lottery rule, and existence of reserved seats. TTC is the most efficient mechanism, followed by BM and DA. In terms of fairness, TTC has a lot of justified envy using the standard envy measures but looks more favorable in terms of fairness according to other envy measures that I study. Having a constrained list does not seem to play a big role in efficiency but does increase justified envy. Having reserved seats has mixed effects on efficiency but less justified envy. In contrast, single versus multiple tie-breaking lotteries do not seem to differ much in terms of either efficiency or fairness. Finally, sophisticated students gain under a manipulable mechanism at the expense of sincere students, on average.

In this chapter, I normalize the utility of being assigned to the outside option of all students to zero. While this is commonly done in many other cases, this assumption is especially problematic with assignment to a magnet school because students are assigned to a base school prior to the magnet application. Many students attend their base school if their magnet application is unsuccessful, while others exit the school district. Ignoring these heterogeneous options as students' outside options might introduce several issues. Further, the current estimation results suggest that the variance-covariance matrix can introduce identification challenges when the number of schools is large. There might be econometric improvements available by placing restrictions on the variance-covariance matrix, specifically by leveraging work from econometric theory on high dimensional problems. The treatment of the outside option and econometric improvements to the estimation procedure in high dimensional space are two areas of work currently in progress using these data.

Table 2.1 Summary statistics for key demographic variables in WCPSS, 2014

	NGE	AG	Female	White	Asian	Black	Other	Hisp	LEP	SPED
Population	79.73%	16.81%	48.87%	49.00%	7.04%	23.90%	4.42%	15.64%	13.25%	12.33%
Sample	24.42%	27.65%	50.09%	44.47%	18.1%	25.27%	4.92%	7.24%	7.84%	7.22%

Notes: The population includes all students in WCPSS in 2014, 153,650 students in total. The sample includes magnet applicants meeting our sample restrictions, 4,999 students. The sample restrictions exclude applicants for the following reasons: students who withdrew their applications (5 students), students admitted to an early college (112 students), students whose assignment was the result of an administrative assignment (304 students), students with missing information on their lottery number (130 students), and students with missing information on their address (8 students). The variables are defined as follows: NGE=1 for non-grade-entry students, AG=1 for academically gifted students, Female=1 for female students, White=1 for white students, Asian=1 for Asian students, Black=1 for black students, Other=1 for students of a race other than Asian, black, or white, Hisp=1 for Hispanic students, LEP=1 for limited-English proficiency students, and SPED=1 for special education students.

Table 2.2 Number of schools listed (%)

	Num of St	1	2	3
All Applicants	4999	56.39%	19.64%	23.96%
Elementary School Applicants	2026	46.50%	19.30%	34.21%
Middle School Applicants	1748	55.61%	23.74%	20.65%
High School Applicants	1225	73.88%	14.37%	11.76%

Notes: The percentages of students with submitted rank order lists of each length are shown.

Table 2.3 Students with highest four priorities in top ranked school (%), grade-entry applicants

	P1	P2	P3	P4
Elementary School Students	19.92%	NA	NA	NA
Middle School Students	7.2%	14.01%	18.02%	3.63%
High School Students	12.43%	42.43%	9.72%	8.11%

Notes: The percentages of students with each priority level are shown for the highest four priority levels. The highest priority levels are siblings (P1), for secondary schools only, magnet pathway (P2), magnet non-pathway (P3), and non-magnet pathway (P4). The second and third highest priority levels for secondary schools are reserved for students who currently hold magnet seats and are applying to the magnet middle or high school that follows their magnet program's pathway/feeder schools (second highest) or are applying to a magnet school that is not on their pathway (third highest). The fourth highest priority level for secondary schools is reserved for base students at magnet schools who are applying to the magnet middle or high school that follows their magnet program's pathway/feeder schools. Only grade-entry students are shown.

Table 2.4 Elementary school grade-entry cutoffs and assignment rates

School	Cutoffs	Assignrate1	Assignrate2	Assignrate3
Brentwood E	20000.06	1.00	0.00	0.00
Brooks E	20066.50	0.25	0.00	0.00
Bugg E	20000.72	0.78	0.00	0.00
Combs E	20016.67	0.43	0.00	0.00
Conn E	20000.58	0.82	0.00	0.00
Douglas E	20068.63	0.30	0.00	0.00
Farmington E	20066.83	0.42	0.00	0.00
Fox Road E	20002.28	1.00	0.00	0.00
Fuller E	20004.82	0.65	0.00	0.00
Green E	10000.94	1.00	0.21	0.00
Hunter E	20000.66	0.80	0.00	0.00
Joyner E	20008.77	0.42	0.00	0.00
Kingswood E	20004.91	0.44	0.00	0.00
Millbrook E	20000.21	0.95	0.00	0.00
Partnership E	10004.70	1.00	0.67	0.00
Poe E	20000.70	0.64	0.00	0.00
Powell E	20000.32	0.89	0.00	0.00
Smith E	10004.58	1.00	1.00	0.00
Underwood E	20004.43	0.69	0.00	0.00
Washington E	20066.81	0.46	0.00	0.00
Wendell E	20002.13	0.65	0.00	0.00
Wiley E	20066.54	0.25	0.00	0.00
Zebulon E	20002.65	0.80	0.00	0.00

Notes: Cutoffs are the lowest eligibility score among all students who are assigned to the school, expressed in terms of the point value for the priority received by a student at a school. Priority points are defined by sum the points that depend on where you rank the school, the points that depend on the priority structure, and the tie-breaking lottery number. Points for the rank order are 20,000 for the top ranked school, 10,000 points for the second ranked school, and 0 for the third ranked school. Points from priorities range from 0 to 256 (1,024) for elementary (secondary) school applicants. The lottery numbers are uniformly distributed on (0,1). Assignrate1, Assignrate2, and Assignrate3 are the assignment rates for students who ranked the school first, second, and third, respectively. That is, these are the percentages of students assigned to the school by position on the submitted list. Only grade-entry students are shown.

Table 2.5 Middle school grade-entry cutoffs and assignment rates

School	Cutoffs	Assignrate1	Assignrate2	Assignrate3
Carnage M	20256.42	0.40	0.00	0.00
Carroll M	20004.16	0.61	0.00	0.00
Centennial M	20064.33	0.50	0.00	0.00
E Garner M	0.00	1.00	1.00	1.00
E Millbrook M	0.00	1.00	1.00	1.00
Ligon M	20256.94	0.34	0.00	0.00
Martin M	20066.74	0.40	0.00	0.00
Moore Square M	20064.35	0.49	0.03	0.03
Zebulon M	20000.40	0.86	0.00	0.00
Carnage M AG	0.00	1.00	1.00	1.00
Ligon M AG	0.00	1.00	1.00	1.00

Notes: See Table 2.4 for definitions. Two schools (Carnage and Ligon) are shown twice: regular program and academically gifted (AG) program. The AG program at these schools is a form of reserved seats for students with high levels of academic achievement. Moore Square Middle has non-zero assignment probabilities for second and third ranking students because of a priority at this school for AG students, as explained in the text. Only grade-entry students are shown.

Table 2.6 High school grade-entry cutoffs and assignment rates

School	Cutoffs	Assignrate1	Assignrate2	Assignrate3
Enloe H	20512.07	0.72	0.00	0.00
Garner H	10004.28	1.00	0.42	0.00
Millbrook H	20128.47	0.40	0.00	0.00
SE Raleigh H	0.00	1.00	1.00	1.00

Notes: See Table 2.4 for definitions. Only grade-entry students are shown.

Table 2.7 Estimation of elementary schools, homogeneous sophistication, individual characteristics

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Brentwood E	2.83	1.02	1.80	-1.71	4.04	2.37	0.76	1.50	-1.39	0.71	6.21
SD	0.76	0.58	1.53	0.57	0.82	0.82	1.03	0.99	1.16	0.98	1.52
Brooks E	15.78	-3.03	-3.53	-0.27	1.90	-1.39	0.78	-0.33	-2.61	1.47	5.30
SD	0.35	0.46	1.33	0.39	0.67	0.78	0.95	1.00	1.05	0.79	1.33
Bugg E	7.04	-0.12	-2.20	-0.36	1.91	2.21	0.07	0.39	0.97	0.35	5.01
SD	0.39	0.41	2.74	0.52	0.74	0.57	0.83	0.87	0.92	0.97	1.50
Combs E	14.47	-1.17	1.84	-0.70	1.42	0.42	3.23	1.05	-3.06	2.10	5.67
SD	0.60	0.47	1.65	0.55	0.55	0.73	1.03	0.90	1.13	0.94	1.87
Conn E	7.56	0.42	1.90	-0.87	3.00	2.40	1.18	0.49	-1.55	1.51	5.21
SD	0.46	0.44	1.27	0.42	0.58	0.60	0.87	0.84	0.93	1.00	1.48
Douglas E	13.79	-6.61	1.13	0.35	0.57	1.51	2.36	1.26	-2.05	2.49	4.13
SD	0.43	0.47	1.25	0.32	0.71	0.56	1.02	0.71	0.89	0.86	0.90
Farmington E	13.60	-5.25	-4.94	-0.49	2.28	0.39	1.22	-2.93	-1.45	-2.53	4.11
SD	0.42	0.33	0.91	0.60	0.69	0.63	1.31	0.77	0.87	1.18	1.10
Fox Road E	4.07	-2.35	-5.69	-1.20	7.31	2.67	1.80	0.78	-3.55	2.20	6.36
SD	0.79	0.46	1.29	0.52	0.83	1.26	1.05	1.04	1.14	1.02	1.42
Fuller E	10.16	0.98	3.50	-0.45	12.69	0.90	5.38	0.89	-1.09	-0.68	5.17
SD	0.53	0.45	1.61	0.38	0.70	0.65	0.95	0.79	0.84	1.14	1.61
Green E	0.00	2.35	1.44	-1.36	3.78	0.40	-3.39	-1.86	-4.05	0.57	6.97
SD	0.91	0.64	1.80	0.55	0.89	0.97	1.86	1.12	1.25	1.35	1.51
Hunter E	6.57	0.38	4.31	0.05	3.07	1.47	2.55	0.68	-1.24	-1.00	5.38
SD	0.54	0.43	1.25	0.38	0.61	0.54	0.90	0.80	0.82	1.01	1.88
Joyner E	12.67	-2.52	1.71	-0.56	3.05	1.54	3.02	4.18	-1.79	1.76	5.01
SD	0.45	0.46	1.74	0.49	0.64	0.62	1.15	0.87	1.12	1.14	1.39
Kingswood E	6.80	-5.40	3.20	0.34	-5.65	0.77	-0.66	3.79	-2.23	5.44	7.09
SD	0.81	0.77	1.98	0.87	1.09	1.48	1.96	2.02	1.65	1.31	2.17
Millbrook E	5.14	-1.06	-3.28	-1.36	5.48	3.06	-0.71	3.10	-1.38	1.04	5.13
SD	0.46	0.79	1.41	0.41	0.85	0.71	1.25	0.67	0.96	0.90	1.21

Tables 2.7 (continued)

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Partnership E	2.09	0.99	-1.26	-1.06	-0.74	1.65	1.50	0.92	-3.39	2.50	5.53
SD	0.67	0.54	1.54	0.51	0.74	0.61	1.18	0.85	1.28	1.37	1.64
Poe E	4.27	-2.20	3.04	-0.32	4.25	2.81	0.96	1.31	-2.03	1.26	6.32
SD	0.67	0.45	1.63	0.44	0.62	0.71	1.12	0.97	1.07	1.23	2.17
Powell E	4.03	-1.00	1.66	-0.12	3.67	1.24	0.55	0.33	1.24	-1.17	4.82
SD	0.54	0.52	1.43	0.49	0.90	0.57	1.10	0.86	1.13	1.09	1.59
Smith E	4.33	-0.57	-2.07	0.06	0.73	-0.45	-1.31	-0.95	-2.41	1.78	4.52
SD	0.50	0.52	3.29	0.53	0.72	0.84	1.07	0.84	0.88	0.84	0.95
Underwood E	8.97	0.69	-1.20	-0.37	5.57	-0.94	0.90	-3.31	-3.13	0.36	4.54
SD	0.40	0.38	1.85	0.51	0.58	0.74	1.21	2.36	1.25	1.03	1.17
Washington E	16.06	-2.63	1.37	-1.01	9.56	0.08	3.49	0.25	-0.32	2.55	4.51
SD	0.44	0.27	1.16	0.38	0.44	0.39	0.72	0.73	0.82	0.60	1.02
Wendell E	12.42	1.70	-0.65	-0.10	6.59	-0.57	-1.82	-5.06	2.18	-1.81	6.39
SD	1.03	0.72	1.29	0.72	0.58	0.64	1.28	1.37	1.13	1.39	1.39
Wiley E	16.55	-2.84	1.81	-0.56	4.06	-0.15	3.66	-0.86	-0.27	1.61	4.10
SD	0.40	0.30	1.23	0.33	0.80	0.47	0.73	0.79	0.63	0.81	1.13
Zebulon E	12.84	-2.72	-0.83	0.29	9.34	1.35	-4.61	-1.52	0.31	0.52	7.02
SD	0.64	0.90	1.96	0.37	0.52	0.44	1.18	0.78	1.29	0.90	1.58

Notes: See Table 2.1 for definitions. The intercept is interpreted as a school fixed effect. Unobs is the standard deviation of unobserved characteristics and reflects the variation across students in preferences for the school. The first row for each school provides the means of the posterior distribution, while the second row provides its standard deviation.

Table 2.8 Estimation of middle schools, homogeneous sophistication, individual characteristics

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Carnage M	16.68	2.10	-3.16	-0.04	4.91	1.41	2.67	-0.80	1.22	-1.07	5.70
SD	0.43	0.62	0.40	0.34	0.52	0.52	0.86	0.86	0.72	0.73	1.20
Carroll M	13.07	-1.87	-5.42	-0.67	0.62	1.30	2.30	0.54	1.40	-0.39	6.42
SD	0.59	0.71	0.74	0.44	0.89	0.72	1.10	1.03	1.02	0.82	1.52
Centennial M	15.44	-2.03	-6.20	-0.72	-0.50	2.16	3.03	-0.17	-0.12	0.59	9.08
SD	0.70	0.81	0.70	0.57	0.94	0.80	1.30	1.31	1.14	0.99	2.40
E Garner M	3.74	6.37	-1.42	2.40	4.98	2.38	4.07	0.86	1.36	-0.68	10.99
SD	0.75	1.20	0.81	0.78	1.27	0.87	2.01	1.62	1.52	1.21	2.55
E Millbrook M	1.75	5.08	-4.47	2.75	4.30	5.32	4.77	1.31	-0.89	0.24	12.38
SD	1.01	1.17	0.96	0.80	1.21	0.97	2.01	1.91	1.98	1.20	2.97
Ligon M	17.24	1.41	-3.52	-0.13	2.87	-0.35	1.65	-0.52	-0.17	-1.72	5.20
SD	0.41	0.58	0.41	0.31	0.50	0.61	0.85	0.76	0.68	0.67	1.41
Martin M	18.13	2.36	-5.63	-0.37	1.47	-1.33	0.40	-1.63	1.40	-1.77	5.62
SD	0.44	0.72	0.41	0.39	0.50	0.72	0.83	0.89	0.65	0.80	1.26
Moore Square M	13.66	1.83	-4.47	-0.12	1.69	1.18	2.55	1.20	0.22	1.34	8.04
SD	0.56	0.77	0.50	0.49	0.76	0.59	1.12	1.06	1.05	0.89	2.20
Zebulon M	8.58	-6.78	-2.42	2.85	4.56	4.41	7.07	-0.21	3.75	0.74	7.89
SD	0.59	1.34	0.66	0.52	1.11	0.66	1.41	1.33	1.40	0.94	2.19

Notes: See Table 2.7 for definitions.

Table 2.9 Estimation of high schools, homogeneous sophistication, individual characteristics

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Enloe H	21.95	-6.50	3.26	1.41	6.71	-3.12	3.58	-2.23	0.44	-1.15	6.40
SD	0.73	0.95	0.84	0.62	1.40	0.70	1.80	1.24	1.27	1.48	2.00
Garner H	13.06	-4.12	0.20	-1.56	0.34	3.80	0.82	0.51	1.40	-0.24	7.73
SD	0.71	0.88	0.71	0.55	0.96	0.74	1.82	1.08	1.01	1.19	2.04
Millbrook H	17.76	-4.75	-3.35	0.63	-0.99	1.97	4.27	-1.70	0.81	-2.87	4.94
SD	0.61	0.72	0.58	0.57	1.27	0.54	1.52	0.81	0.82	0.82	1.37
SE Raleigh H	4.80	6.22	-3.48	-2.23	4.68	7.96	4.85	4.95	-1.44	2.20	17.59
SD	1.47	2.05	1.65	1.19	2.21	1.49	3.16	2.30	2.08	2.49	4.17

Notes: See Table 2.7 for definitions.

Table 2.10 Estimation of elementary schools, homogeneous sophistication, school characteristics

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Brentwood E	-5.76	1.49	2.02	-1.61	4.30	2.89	1.47	2.47	-0.72	-0.32	9.31
SD	2.40	0.71	2.10	0.57	0.90	0.73	1.27	1.05	1.27	1.29	2.98
Brooks E	6.14	-3.73	-1.27	-0.13	0.36	-1.04	2.60	-0.51	-1.06	0.99	6.26
SD	3.29	0.48	2.29	0.48	0.96	0.59	1.07	0.96	1.67	1.15	2.10
Bugg E	0.04	-0.13	0.89	0.41	2.72	2.90	-0.23	0.48	0.61	-0.29	6.45
SD	2.46	0.54	1.72	0.49	0.91	0.75	1.32	0.96	0.98	0.96	1.74
Combs E	5.14	-1.13	3.92	-0.78	0.63	0.79	4.79	1.60	-2.10	2.14	7.55
SD	2.52	0.51	2.44	0.51	0.95	0.67	1.16	0.93	1.21	1.21	2.28
Conn E	-1.49	0.93	1.93	-0.43	2.37	3.48	1.99	0.78	-1.94	1.02	7.38
SD	2.20	0.53	1.30	0.41	0.69	0.60	1.08	0.95	0.98	0.97	2.09
Douglas E	7.52	-8.04	3.58	0.03	0.37	0.77	2.89	1.30	-0.51	0.73	5.16
SD	3.20	0.62	2.01	0.44	0.75	0.76	1.23	0.94	1.08	1.19	2.10
Farmington E	4.45	-6.00	-3.63	-0.41	2.55	1.68	-0.16	-0.67	-0.46	-2.42	4.07
SD	3.72	0.48	1.50	0.48	0.53	0.79	1.10	1.28	1.06	1.04	1.50
Fox Road E	-4.54	-2.42	-8.54	-1.24	6.95	2.72	1.32	2.11	-3.57	1.86	8.00
SD	4.05	0.99	2.10	0.61	1.01	0.77	1.47	1.32	1.73	1.51	2.78
Fuller E	-0.34	0.86	5.01	-0.34	14.00	2.07	7.16	1.28	-0.73	-2.34	8.37
SD	2.43	0.70	2.41	0.53	1.01	0.74	1.55	1.20	1.35	1.80	2.03
Green E	-8.10	2.91	-0.33	-1.60	1.76	-0.88	-4.25	-1.07	-3.10	-0.52	11.55
SD	4.29	0.93	3.99	0.79	1.23	1.05	1.75	1.29	1.30	1.57	4.07
Hunter E	-3.39	0.61	5.41	0.31	2.66	2.53	3.39	1.11	-1.14	-2.56	8.21
SD	2.28	0.51	1.65	0.52	0.65	0.65	1.25	1.15	1.02	1.29	2.31
Joyner E	3.84	-3.31	2.15	-0.74	1.07	2.58	4.62	4.79	-1.36	1.39	7.90
SD	2.54	0.63	2.89	0.62	0.89	0.69	1.62	0.94	1.40	1.41	2.55
Kingswood E	-2.48	-4.44	2.21	0.02	-7.01	0.74	-0.43	4.97	-2.78	6.92	11.81
SD	3.33	0.95	3.46	0.92	1.15	1.10	2.90	1.67	1.89	1.62	3.32

Tables 2.10 (continued)

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Millbrook E	-5.78	0.16	-6.73	-0.40	4.69	4.40	1.30	3.77	-0.92	1.48	6.72
SD	2.52	0.71	4.26	0.49	0.74	0.71	1.44	0.85	1.15	1.15	2.02
Partnership E	-8.09	1.65	0.36	-0.18	-1.08	2.52	1.30	0.89	-4.05	2.45	8.84
SD	6.81	0.71	3.01	0.82	0.83	0.86	1.21	1.15	1.62	1.23	2.91
Poe E	-7.30	-2.67	2.68	0.24	4.74	4.81	2.07	3.11	-1.80	0.63	10.30
SD	2.79	0.62	2.03	0.94	1.11	1.10	1.50	1.74	1.16	1.35	3.54
Powell E	-2.73	-0.74	1.03	0.25	2.34	1.64	0.97	1.00	1.65	-1.87	7.02
SD	3.12	0.66	1.81	0.50	0.73	0.75	1.17	1.05	1.44	1.14	2.15
Smith E	-5.04	-1.06	-11.83	0.04	-1.98	0.47	-4.81	0.96	-0.58	0.12	6.12
SD	4.76	0.67	3.52	0.53	1.45	0.63	1.90	1.14	1.25	1.14	1.55
Underwood E	-2.52	0.61	1.10	-0.41	6.36	-0.14	0.85	-0.86	-3.72	-1.23	6.05
SD	3.83	0.58	3.14	0.47	0.61	0.55	1.41	1.09	1.51	1.11	2.32
Washington E	7.12	-3.34	2.35	-1.10	10.37	1.05	5.41	0.43	-0.21	1.54	5.75
SD	2.63	0.51	1.95	0.44	0.57	0.68	1.00	1.01	0.94	0.95	1.69
Wendell E	3.40	1.62	-0.76	0.40	6.23	-1.40	-0.98	-5.67	2.68	-5.68	6.47
SD	2.59	0.85	1.99	0.70	0.88	0.78	1.48	1.29	1.82	1.19	2.22
Wiley E	6.29	-3.27	2.83	-0.27	3.03	0.28	5.12	-1.54	0.80	1.07	4.95
SD	3.85	0.46	2.04	0.42	0.70	0.50	0.79	0.71	0.95	0.90	1.78
Zebulon E	3.73	-4.17	8.47	0.63	6.90	2.83	0.53	-0.47	4.02	0.98	4.93
SD	2.23	1.19	1.81	0.83	1.00	0.87	1.55	0.83	1.48	0.98	1.75

Panel B School-constant estimation results

	TradCal	Turnover	SPG
Estimates	-1.74	0.17	0.13
SD	7.01	0.44	0.11

Notes: See Table 2.7 for definitions. Panel B provides the estimates of the school characteristics. TradCal=1 if the school has a traditional calendar, as opposed to a year-round calendar. Turnover is the percentage of teachers who left WCPSS. SPG is the school performance grade, which is measured on a 100 point scale and is a function of school achievement and academic growth.

Table 2.11 Estimation of middle schools, homogeneous sophistication, school characteristics

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Carnage M	-3.26	2.26	-5.05	-0.31	5.16	1.42	2.35	-0.82	1.60	-1.46	6.77
SD	7.06	0.73	0.58	0.47	0.62	0.73	0.92	1.17	0.87	0.90	1.92
Carroll M	7.54	-2.32	-6.66	-0.90	-0.74	1.79	1.99	0.78	2.55	-0.03	8.71
SD	3.75	1.02	0.86	0.67	1.34	0.85	1.34	1.42	1.21	1.03	2.81
Centennial M	2.63	-3.01	-9.34	-1.26	-1.34	1.68	2.87	-0.98	-0.30	0.59	11.47
SD	5.93	1.11	0.87	0.75	1.11	0.95	1.53	1.71	1.56	1.22	3.39
E Garner M	-6.56	7.84	-5.81	4.43	1.86	5.91	5.30	0.04	2.69	-2.56	18.45
SD	5.73	2.04	1.63	1.41	1.99	1.65	2.50	2.88	2.43	2.05	4.77
E Millbrook M	-8.58	6.53	-10.16	4.14	-3.06	7.72	4.57	2.65	0.77	-0.19	20.01
SD	5.29	2.24	1.56	1.77	2.13	1.61	2.47	2.76	2.30	2.19	5.01
Ligon M	3.88	2.05	-4.59	-0.11	2.78	-0.12	1.11	-0.55	0.31	-1.87	6.36
SD	4.22	0.70	0.66	0.46	0.64	0.70	0.93	1.10	0.92	0.91	1.81
Martin M	0.89	2.38	-7.17	-0.24	1.13	-1.80	-0.68	-1.83	1.57	-1.71	6.57
SD	5.19	0.77	0.57	0.44	0.61	0.62	1.06	1.11	0.86	0.90	1.66
Moore Square M	4.18	1.23	-6.33	0.19	0.82	1.33	2.34	1.09	0.24	1.77	9.79
SD	6.41	0.94	0.74	0.63	0.99	0.94	1.23	1.38	1.35	1.11	2.40
Zebulon M	-2.05	-4.85	-4.35	2.65	-3.33	1.70	6.21	-3.75	6.04	1.14	7.10
SD	4.38	1.41	1.44	0.92	1.44	1.43	1.79	2.52	1.90	1.72	2.51

Panel B School-constant estimation results

	TradCal	Turnover	SPG
Estimates	-8.08	-0.67	0.50
SD	6.50	0.52	0.18

Notes: See Tables 2.7 and 2.10 for definitions.

Table 2.12 Estimation of high schools, homogeneous sophistication, school characteristics

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Enloe H	7.31	-6.54	4.24	1.29	8.16	-3.63	2.88	-2.59	0.43	-0.44	8.32
SD	5.38	1.46	1.02	0.78	1.68	1.10	2.20	1.70	1.51	1.74	3.13
Garner H	-4.79	-4.35	0.97	-1.83	1.08	4.21	0.04	0.69	1.25	0.52	9.53
SD	8.90	1.29	0.97	0.74	1.98	1.01	1.74	1.49	1.18	1.35	4.12
Millbrook H	3.30	-4.19	-3.30	0.35	-1.21	1.40	3.62	-2.50	1.07	-1.89	5.66
SD	6.84	1.10	0.68	0.61	1.25	0.93	1.67	1.38	1.32	1.39	2.20
SE Raleigh H	-6.94	5.88	-2.38	-2.91	6.48	8.83	4.25	5.23	-2.52	3.16	19.89
SD	6.95	2.57	1.84	1.53	2.76	1.87	3.25	2.89	2.48	2.65	7.56

Panel B School-constant estimation results

	Turnover	SPG
Estimates	-0.40	0.29
SD	0.99	0.21

Notes: See Tables 2.7 and 2.10 for definitions.

Table 2.13 Estimation of middle schools, heterogeneous sophistication

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Carnage M	1.33	-1.97	8.94	-3.57	18.53	11.78	7.11	1.20	4.87	-2.26	273.35
SD	7.28	2.61	2.33	2.24	2.53	2.56	2.81	2.81	2.67	2.69	169.70
Carroll M	-29.45	-6.46	-15.84	-12.35	-5.86	-4.24	-0.77	-0.90	-1.80	-1.38	1178.80
SD	12.06	9.98	10.28	9.66	9.80	9.60	9.86	9.85	9.75	9.75	888.20
Centennial M	5.72	-4.40	-9.10	-3.36	-3.07	0.81	1.76	-2.86	-4.31	3.11	273.17
SD	6.68	2.59	2.97	2.27	2.49	2.39	2.92	2.79	2.85	2.72	169.65
E Garner M	3.19	8.04	10.52	3.04	3.64	-1.33	-0.74	0.72	0.01	-4.71	273.37
SD	6.47	2.61	2.39	2.32	2.49	2.33	2.86	2.64	2.63	2.52	169.82
E Millbrook M	7.63	-2.60	-1.00	1.46	0.39	1.96	0.34	1.66	-4.19	-4.68	461.26
SD	5.36	3.94	3.69	3.64	3.70	3.63	4.09	4.08	3.90	3.84	311.29
Ligon M	-8.49	-3.42	17.72	-2.99	-12.28	-6.38	-2.29	-2.72	-2.95	1.19	274.67
SD	6.01	3.87	3.18	3.07	3.51	3.83	4.10	4.02	3.55	3.25	170.33
Martin M	4.87	0.46	-0.15	1.47	2.57	-15.87	-9.49	-10.26	3.12	-1.76	273.39
SD	5.51	2.98	2.39	2.35	2.57	2.49	3.21	2.96	2.94	2.84	169.75
Moore Square M	2.81	3.68	-8.90	2.01	-0.54	6.36	2.15	4.01	2.72	5.18	785.96
SD	8.08	7.88	7.38	7.29	7.62	7.48	7.96	7.89	7.74	7.67	431.44
Zebulon M	10.10	-6.18	10.91	2.18	6.06	-4.63	-1.76	1.11	1.56	-1.97	274.35
SD	4.72	2.79	2.42	2.34	2.53	2.48	2.78	2.83	2.67	2.77	170.58

Panel B School-constant estimation results

	TradCal	Turnover	SPG
Estimates	-10.55	0.16	0.52
SD	6.97	0.61	0.19

Notes: See Tables 2.7 and 2.10 for definitions.

Table 2.14 Estimation of high schools, heterogeneous sophistication

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Enloe H	7.18	-3.83	5.80	-0.32	4.40	-2.01	1.44	-4.85	3.14	1.49	5.82
SD	5.30	0.82	0.77	0.63	0.84	0.82	1.37	1.17	1.25	1.65	1.69
Garner H	-4.26	-2.08	1.14	-3.15	0.68	4.08	-0.69	0.79	1.71	2.19	13.15
SD	8.86	1.59	1.48	1.06	2.27	1.37	2.36	1.96	1.78	2.14	3.47
Millbrook H	0.27	-0.71	-0.93	-0.81	-3.25	2.83	2.15	-4.88	3.54	-0.20	6.56
SD	6.82	1.28	0.81	0.75	2.03	0.98	1.53	1.14	1.51	1.88	2.38
SE Raleigh H	-3.73	1.19	-1.60	-3.77	3.74	7.37	2.81	1.78	-0.75	3.16	16.95
SD	6.90	1.97	1.61	1.25	1.98	1.50	2.93	2.32	2.25	2.32	4.10

Panel B School-constant estimation results

	Turnover	SPG
Estimates	0.00	0.23
SD	0.99	0.21

Notes: See Tables 2.7 and 2.10 for definitions.

Table 2.15 Ranking behaviors summary statistics

Panel A Individual Behaviors

	Elementary	Middle	High
Truth-telling	1.68%	7.89%	11.18%
Truncated truth-telling	43.18%	61.90%	72.08%
Skip-the-top	15.72%	4.23%	6.53%
Skip-the-middle	20.65%	15.68%	4.24%
Flipping in position 1	7.32%	2.06%	1.80%
Flipping in position 2	11.44%	8.24%	4.16%

Panel B Grouping Behaviors

	Elementary	Middle	High
Flipping	18.77%	10.30%	5.96%
Ranking 1st school truthfully	76.95%	93.71%	91.66%

Notes: The strategies in Panel A are mutually exclusive and exhaustive. The strategies in Panel B group together related strategies from Panel A. The six strategies are truth-telling, truncated truth-telling, skip-the-top, skip-the-middle, flipping starting in position 1 (flip1), and flipping starting in position 2 (flip2). Truth-telling means a student reports a ranking that is exactly the same as their true preference, up to the constraint of the mechanism. Truncated truth-telling means a student reports truthfully but reports a truncated version of their true preferences. Skip-the-top means a student does not report her true favorite school first and instead reports a school somewhere in the middle of her true preferences. Skip-the-middle means a student reports her true favorite school first but then omits her next truly preferred school. Flip1 means a student does not report her true favorite school first and the second school she ranks is in fact more preferred to the first ranked school. Flip2 means a student reports the first school truthfully, but the second ranked school is in fact lower in the true preference list than the third ranked school. In Panel B, flipping aggregates flip1 and flip2. Ranking 1st school truthfully includes any report with the true favorite reported first.

Table 2.16 Elementary school efficiency comparison, all students

	Single Tie Breaking			Multiple Tie Breaking		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	4.17	20.59%	29.58%	4.17	20.62%	29.58%
constrained BM	4.11	19.98%	29.31%	4.12	20.10%	29.44%
unconstrained BM	4.18	19.84%	29.65%	4.20	20.08%	29.74%
constrained DA	4.07	19.53%	28.70%	4.04	19.31%	28.56%
unconstrained DA	4.16	19.64%	29.18%	4.13	19.28%	28.96%

Notes: Efficiency is measured cardinally by Utility (the average utility) and ordinally by Rate1 and Rate3 (the percentage of students assigned to the true favorite or one of their three favorite schools). The mechanisms considered are TTC, constrained BM, unconstrained BM, constrained DA, and unconstrained DA. Constrained versions of the mechanisms restrict students to submitting at most three schools. Unconstrained BM in elementary school is in fact constrained to submitting at most eleven schools; this is done for computational reasons as explained in the text. Single tie breaking runs one lottery for all schools, while multiple tie breaking runs a separate lottery for each school.

Table 2.17 Middle school efficiency comparison, no reserved seats

Panel A All Students

	Single Tie Breaking			Multiple Tie Breaking		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	30.64	40.68%	57.62%	30.64	40.81%	57.62%
constrained BM	30.12	39.27%	57.25%	30.10	39.38%	57.30%
unconstrained BM	30.13	39.19%	57.27%	30.09	39.31%	57.36%
constrained DA	29.78	36.92%	57.19%	29.58	35.89%	57.05%
unconstrained DA	29.68	36.61%	56.64%	29.41	35.18%	56.16%

Panel B Sophisticated Students

	Single Tie Breaking			Multiple Tie Breaking		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	30.93	43.36%	59.41%	31.01	43.57%	59.55%
constrained BM	30.54	40.67%	59.84%	30.57	40.82%	59.91%
unconstrained BM	30.68	40.92%	60.34%	30.69	41.06%	60.41%
constrained DA	29.96	38.85%	58.29%	29.83	37.90%	58.12%
unconstrained DA	29.85	38.63%	58.00%	29.65	37.34%	57.57%

Panel C Sincere Students

	Single Tie Breaking			Multiple Tie Breaking		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	28.73	22.80%	45.66%	28.17	22.40%	44.75%
constrained BM	27.37	29.93%	39.99%	26.96	29.82%	39.92%
unconstrained BM	26.49	27.66%	36.86%	26.12	27.66%	36.99%
constrained DA	28.53	24.03%	49.87%	27.90	22.48%	49.90%
unconstrained DA	28.56	23.15%	47.55%	27.81	20.80%	46.71%

Notes: See Table 2.16 for definitions. Panels B and C separate students according to their strategic types, sophisticated or sincere. Each mechanism in this table does not include any reserved seats.

Table 2.18 Middle school efficiency comparison for reserved seats, all students

Panel A Single Tie Breaking

	No Reserved Seats			Reserved Seats		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	30.64	40.68%	57.62%	28.82	55.33%	65.72%
constrained BM	30.12	39.27%	57.25%	28.59	52.85%	64.57%
unconstrained BM	30.13	39.19%	57.27%	28.87	52.95%	64.78%
constrained DA	29.78	36.92%	57.19%	27.75	49.50%	64.05%
unconstrained DA	29.68	36.61%	56.64%	27.80	49.11%	63.39%

Panel B Multiple Tie Breaking

	No Reserved Seats			Reserved Seats		
Mechanism	Utility	Rate1	Rate3	Utility	Rate1	Rate3
TTC	30.64	40.81%	57.62%	28.75	55.39%	65.44%
constrained BM	30.10	39.38%	57.30%	28.66	52.92%	64.69%
unconstrained BM	30.09	39.31%	57.36%	28.88	53.14%	64.93%
constrained DA	29.58	35.89%	57.05%	27.70	49.21%	63.90%
unconstrained DA	29.41	35.18%	56.16%	27.71	48.62%	63.06%

Notes: See Table 2.16 for definitions. Each mechanism in this table is shown with and without reserved seats.

Table 2.19 High school efficiency comparison

Panel A All Students

Mechanism	Single Tie Breaking Utility	Rate1	Rate3	Multiple Tie Breaking Utility	Rate1	Rate3
TTC	13.38	61.51%	75.54%	13.37	61.49%	75.54%
constrained BM	13.76	62.76%	75.76%	13.75	62.63%	75.74%
unconstrained BM	13.79	62.92%	75.75%	13.77	62.78%	75.76%
constrained DA	13.45	61.32%	75.76%	13.42	61.03%	75.72%
unconstrained DA	13.46	61.74%	75.63%	13.42	61.38%	75.56%

Panel B Sophisticated Students

Mechanism	Single Tie Breaking Utility	Rate1	Rate3	Multiple Tie Breaking Utility	Rate1	Rate3
TTC	14.42	67.59%	78.54%	14.41	67.58%	78.50%
constrained BM	14.79	67.75%	80.58%	14.77	67.58%	80.47%
unconstrained BM	14.83	67.89%	80.57%	14.80	67.75%	80.54%
constrained DA	14.41	66.39%	78.73%	14.37	66.09%	78.65%
unconstrained DA	14.42	66.88%	78.74%	14.38	66.52%	78.65%

Panel C Sincere Students

Mechanism	Single Tie Breaking Utility	Rate1	Rate3	Multiple Tie Breaking Utility	Rate1	Rate3
TTC	7.37	26.45%	58.28%	7.39	26.39%	58.51%
constrained BM	7.86	34.02%	47.93%	7.92	34.08%	48.48%
unconstrained BM	7.81	34.21%	47.96%	7.85	34.10%	48.17%
constrained DA	7.88	32.08%	58.67%	7.88	31.87%	58.87%
unconstrained DA	7.91	32.07%	57.64%	7.91	31.73%	57.76%

Notes: See Table 2.16 for definitions.

Table 2.20 Elementary school fairness comparison

Panel A Single Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	6834.57	524.07	26.66%	13.11	62.05	3.16%
constrained BM	906.35	55.39	2.82%	16.26	271.42	13.81%
unconstrained BM	2767.72	159.09	8.09%	19.49	394.98	20.09%
constrained DA	556.99	23.00	1.17%	23.18	206.07	10.48%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Panel B Multiple Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	6822.90	524.33	26.67%	13.08	62.06	3.16%
constrained BM	963.62	54.64	2.78%	17.48	268.13	13.64%
unconstrained BM	2704.21	176.12	8.96%	18.15	385.54	19.61%
constrained DA	597.76	23.16	1.18%	24.52	209.65	10.66%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Notes: Fairness of each mechanism is shown, measured by the presence of justified envy. Envy is the number of instances of justified envy. Jealous St is the number of students who have at least one instance of justified envy (also shown as a percentage of all students, % Jealous). Envy Per Jealous St is Envy divided by Jealous St. Envied St is the number of students who are envied by at least one other student (also shown as a percentage of all students, % Envied).

Table 2.21 Middle school fairness comparison, no reserved seats

Panel A Single Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	23134.03	489.71	28.02%	47.39	226.59	12.96%
constrained BM	3417.94	140.11	8.02%	27.78	499.81	28.59%
unconstrained BM	3293.21	108.91	6.23%	31.93	462.22	26.44%
constrained DA	958.94	25.23	1.44%	42.41	304.18	17.40%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Panel B Multiple Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	23087.71	485.05	27.75%	47.82	224.32	12.83%
constrained BM	3442.88	141.83	8.11%	27.14	488.33	27.94%
unconstrained BM	3324.82	108.69	6.22%	32.15	452.27	25.87%
constrained DA	842.03	22.34	1.28%	42.07	261.47	14.96%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Notes: See Table 2.20 for definitions. Each mechanism in this table does not include any reserved seats.

Table 2.22 Middle school fairness comparison, reserved seats

Panel A Single Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	10879.17	274.10	15.68%	40.07	200.85	11.49%
constrained BM	4215.99	102.38	5.86%	42.63	621.95	35.58%
unconstrained BM	3823.82	95.75	5.48%	40.93	565.05	32.33%
constrained DA	1285.61	25.56	1.46%	53.92	431.94	24.71%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Panel B Multiple Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	11166.14	271.74	15.55%	41.46	195.58	11.19%
constrained BM	4011.58	99.25	5.68%	42.08	597.99	34.21%
unconstrained BM	3770.25	100.12	5.73%	38.83	518.03	29.64%
constrained DA	947.26	21.60	1.24%	45.97	330.19	18.89%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Notes: See Table 2.20 for definitions. Each mechanism in this table includes reserved seats.

Table 2.23 High school fairness comparison

Panel A Single Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	2717.38	276.15	22.54%	9.82	43.23	3.53%
constrained BM	2326.77	29.61	2.42%	76.39	279.62	22.83%
unconstrained BM	1686.25	29.35	2.40%	56.86	249.46	20.36%
constrained DA	583.65	6.31	0.51%	83.24	207.32	16.92%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Panel B Multiple Tie Breaking

Mechanism	Envy	Jealous St	% Jealous	Envy Per Jealous St	Envied St	% Envied
TTC	2744.95	275.96	22.53%	9.92	42.99	3.51%
constrained BM	2370.43	31.81	2.60%	72.34	278.70	22.75%
unconstrained BM	1880.06	31.03	2.53%	60.66	265.66	21.69%
constrained DA	594.12	6.14	0.50%	89.00	209.62	17.11%
unconstrained DA	0.00	0.00	0.00%	0.00	0.00	0.00%

Notes: See Table 2.20 for definitions.

CHAPTER

3

DEMAND ESTIMATION WITH HIGH-DIMENSIONAL COVARIANCE MATRIX USING BAYESIAN GRAPHICAL LASSO ESTIMATION

3.1 Introduction

Multivariate probit (MVP) models are good candidates for many topics in empirical research, especially in structural demand estimations. With a random utility model, MVP models allow the errors to be correlated and heterogeneous, and therefore they are a better fit than logit models in many cases, such as school choice and residential choice. However, the application of demand estimation with the MVP model is restricted by the computational difficulties related to the covariance matrix estimation. The number of parameters in the covariance matrix grows quadratically in the market size. Taking the school demand estimation in Wake County Public School System (WCPSS) as an example, there are over five hundred parameters to estimate in the covariance matrix for elementary schools alone and the data contain slightly more than two thousand observations. With the traditional estimation method described in Chapter 2, the elementary school market shows very weak convergence under the homogeneous sophistication model, and it fails to converge under the heterogeneous sophistication model. Even outside the MVP context, estimating a covariance matrix from high dimensional data has been a challenge, and yet it is a difficulty we often face in practice. Gaussian graphical models (GGM) are often used to study the underlying relationships among multivariate variables (that is, the covariance matrix) that follow a Gaussian distribution [Lau96]. Certain non-Gaussian graphical models have been studied as well; however, they are not discussed here, because this paper focuses on demand estimation of MVP models. In this chapter, the term *graphical model* refers to *Gaussian graphical model*.

There two types of graphical models that are widely used in practice: undirected graphs and bi-directed graphs. For Gaussian graphical model, undirected graph is used to model the inverse of covariance matrix, which is also called as precision matrix or concentration matrix; bi-directed graph is used to model covariance matrix. The study of high dimensional graphical model has been a popular topic. Classical methods often rely on penalized likelihood. The most common penalized method is known as the graphical lasso; for example, see [MB06]; [YL07]; [Fri08]; [Yua08]; [Guo11]; [Wit11]. Other penalized approaches include the adaptive graphical lasso and the smoothly clipped absolute deviation (SCAD) penalty developed by [Fan09], and the sparse permutation invariant covariance estimator (SPICE) developed by [Rot08]. Another popular approach in the classical framework is the thresholding method ([BL08]; [Rot09]; [CL11]), but it is hard to satisfy the positive definite constraint under thresholding method. There are also an increasing number of papers studying high dimensional graphical model under the Bayesian framework, such as [Dob11]; [WL12]; [Wan12]; [Wan15]; [BG14]; [BG15]; [MW15].

In this chapter, I adopt the recently developed Bayesian graphical Lasso model to MVP model estimation. I adapt [Wan12]'s Bayesian adaptive graphical Lasso by building a new Gibbs Sampling scheme to estimate MVP models with a penalized covariance matrix. The methodology allows estimation in school systems with up to hundreds of schools, which is an important improvement relative to the proceeding model that has computational difficulties above around 20 schools. In this chapter, I show that the combining the MVP demand estimation with Bayesian graphical lasso can solve the covariance matrix problem without imposing any additional computation burden. The estimation results with the new method converge well and the estimates are well behaved.

The rest of the chapter is organized as follows: Section 2 summarizes related literature and explains why I choose the method of [Wan12], Section 3 illustrates the estimation of MVP models with Bayesian graphical lasso and shows the estimation results, Section 4 discusses the Bayesian structure learning potentials, and Section 5 concludes.

3.2 Background on graphical models

In this section, I provide notation and background on graphical models. To be consistent with literature on graphical models, this section uses separate notation as the rest of the paper. Let

$y = (y_1, y_2, \ldots, y_p)'$ be the p-dimensional random vector from a multivariate normal distribution with zero mean and covariance matrix Σ, $N(0, \Sigma)$. Denote the precision matrix by Ω, and $\Omega = \Sigma^{-1}$. Further, define $\Sigma \equiv (\sigma_{ij})$ and $\Omega \equiv (\omega_{ij})$, where σ_{ij} is the (i, j)th element of Σ and ω_{ij} is the (i, j)th element of Ω.

3.2.1 Precision matrix and undirected graphical models

An *undirected graph* is usually denoted as $G = (V, E)$, where V is a non-empty set of nodes and E is the set of edges that represent unordered distinct pairs from V and $E \subseteq \{(i, j) : i < j\}$. Nodes $i, j \in V$ are called *neighbors* if $(i, j) \in E$. Therefore, neighborhood selection (that focuses on the selection around one specific node) is a sub-problem of covariance selection [MB06]. The underlying structure of G is often represented by a set of binary variables $Z = (z_{ij})_{i<j}$, where z_{ij} is the indicator of whether the edge (i, j) exists or not. For Gaussian graphical models, the element in precision matrix is zero whenever there is a missing edge between a pair of corresponding nodes. That means: $z_{ij} = 0 \iff (i, j) \notin E \iff \omega_{ij} = 0 \iff y_i$ is independent of y_j conditional on all remaining variables in y. [Dem72] first conduct covariance selection, by selecting certain zeros in the precision matrix for i.i.d. observations from a multivariate normal distribution. The traditional way of selecting is done by discrete searches of edges to maximize an objective function. An exhaustive search is computationally expensive and can only work with extremely low dimensions. A more commonly used method is the greedy forward search or backward search [Lau96]. In forward searching, the selection first starts with an empty set of edges and then edges are added iteratively. The backward searching works in the opposite direction. The iteration would stop when the maximization stop criterion is reached. When a single edge is being consider in this method, the number of MLE fits is in the quadratic form of the covariance matrix dimension. As a result, this method is also only suitable for low dimensions.

Penalized approaches are introduced to graphical models to solve the dimension problem. The MB algorithm developed by [MB06] is the first to do so. They focus on the neighborhood of a specific node and therefore it is often called the neighborhood selection. They use a L_1-norm penalty on the estimation of neighborhood of each node. θ^i is the coefficient estimates from the lasso regression, indicating the optimal prediction of neighbors around node i, where each estimation is as follows:

$$\hat{\theta}^i = \arg\min_{\theta^i} \left(N^{-1} \| y_i - y\theta^i \|_2^2 + \lambda \| \theta^i \|_1 \right)$$

where $\|\theta^i\|_1 = \sum_{j \in V \setminus \{i\}} |\theta_j^i|$ is the L_1-norm of the coefficient estimates, and λ is the positive shrinkage parameter.

The graphical lasso is developed based on the MB approach and improves the loss function to a regularized likelihood. The graphical lasso imposes penalty on Ω directly and maximizes a convex object function to encourage zeros in Ω, such as [YL07] and [Fri08]. The penalized log-likelihood estimation is

$$\hat{\Omega} = \arg\max_{\Omega}\left(\log(\det\Omega) - \mathrm{tr}\left(\frac{S}{N}\Omega\right) - \lambda\|\Omega\|_1\right)_{M^+}$$

where $S = y'y$, N is the number of observations, λ is the positive shrinkage parameter, $\|\Omega\|_1 = \sum_{1\le i,j\le p} |\omega_{ij}|$ is the L_1-norm of Ω, and the optimization is over the space of positive definite matrices M^+.

In the Bayesian context, traditional methods are based on hierarchical models and impose global conjugate priors that are on the space of positive definite matrices, such as G-Wishart prior. [Gre95] first introduces the reversible jump Markov Chain Monte Carlo (MCMC) and [GG99] future extend it to handle decomposable graphical Gaussian models. Such methods have many advantages, especially the ease of computation from the conjugation. However, it has several restrictions as well. Unrestricted graphical model only works efficiently for decomposable graphs, where the marginal likelihoods are available up to the overall normalizing constants [Rov00], and requires heavy computation for nondecomposable graphs. [Rov00] and [Rov02] therefore introduce hyper inverse Wishart (HIW) priors, which not only handles nondecomposable graphs but also enjoys the easy computation from conjugacy. Unfortunately, the HIW distribution still faces to the challenge from the unknown normalizing constant. Several different computation methods have been proposed to improve the reversible jump MCMC and draw samples from HIW distribution or G-Wishart distribution; for example, see [Jon05]; [CS09]; [WC10]. Particularly, [Dob11] develop the method that's based on the Cholesky decomposition of the precision matrix. [WL12] and [Len13] combine the exchange algorithm with the reversible jump MCMC (double reversible jump) to sample G-Wishart distribution on graphs directly and avoid normalizing constant calculation. But these methods are only suitable for moderate dimension due to the computation burden.[MW15] adapt another trans-dimensional MCMC methodology, birth-death MCMC. They learn the structure of the precision matrix by birth-death algorithm and then sample from the G-Wishart distribution given the structure.

[PC08] develop the Bayesian lasso by showing that the Laplace prior could induce sparsity. However, the Laplace prior is not analytic for Bayesian inference, and [PC08] use a scale mixture of normals in hierarchical models to solve this problem. Following this approach, some recent papers have been trying to combine the lasso estimation with the Bayesian graphical models, such as [Wan12]; [Wan15]; [BG15]. [Wan12] provides a Bayesian interpretation on the graphical lasso and develops the Bayesian graphical lasso, which puts independent exponential prior on diagonal entries, Laplace prior on off-diagonals, and is subject to positive definiteness restriction. The drawback of this method is that it uses continuous priors and cannot produce absolute zeros in the precision (or covariance) matrix. It works well with estimation due to the efficient computation, but it cannot be used to conduct structure leaning simultaneously. As a result, in the case where the number of observation is less than the number of parameters, this method does not work. [Wan15] further improve the method with stochastic search structure learning (SSSL) by using continues spike and slab priors. This method can still estimate precision matrix efficiently and conduct structure learning simultaneously, but it is sensitive to the choice of hyperparameters and shrinkage parameters [Wan15]. [BG15] put point mass at zeros and continuous distribution on off-diagonal entries. This method comes with the fast computation and structure learning, and really works for the high dimension situation. It uses the Laplace approximation, which is based on a Taylor series expansion, instead of the exact posterior distribution. Although they have proved that the approximation error would go to zero asymptotically, it might lead to larger error in the demand estimation where the precision matrix estimation is only one step of the MCMC process.

3.2.2 Covariance matrix and bi-directed graphical models

[CW93] present the covariance matrix Σ using a bi-directed graph $G = (V, E)$, which has a very similar setup as the undirected graph in that V is also the set of modes and E is the set of edges. The main difference is that the edges here are bi-directed, rather than unordered as in the undirected graph for precision matrix. Same as undirected graph, the underlying structure of G can also be represented by a set of binary variables $Z = \left(z_{ij}\right)_{i<j}$. However, the computation is much harder due to the bi-direct, and there are fewer papers that directly estimate covariance matrix comparing to papers that estimate via precision matrix. In the classical framework, [Kau96] develops a maximum likelihood estimator, and it is later extended into different likelihood based estimators such as the Iterative Conditional Fitting by [Cha07]. [Wer06] arrange covariance matrices in covariance chains and derive orthogonal decompositions for covariance chains to ease the computation. However, these methods all come with heavy computation burden. The thresholding method is developed to handle high dimensional estimation by setting an element σ_{ij} to be zero if it is below a certain threshold ([BL08]; [Rot09]; [CL11]). Although the thresholding method is relatively easy to compute, it is hard to guarantee the positive definitive of matrix [Wan15]. On the lasso side, [BT11] develop the covariance graphical lasso by combining lasso and Gaussian covariance model. The penalized log-likelihood estimator is

$$\hat{\Sigma} = \arg\min_{\Sigma}\left(\log(\det\Sigma) - \mathrm{tr}\left(\frac{S}{N}\Sigma^{-1}\right) - \lambda\|\Sigma\|_1\right)_{M^+}$$

where λ is the positive shrinkage parameter, $\|\Sigma\|_1 = \sum_{1\le i,j\le p} |\sigma_{ij}|$ is the L_1-norm of Σ, and the optimization is over the space of positive definite matrices M^+. [Fan09] adapt the adaptive lasso to covariance graphical lasso and develop the SCAD method. The SPICE developed by [Rot08] use penalty on the off-diagonal elements of the covariance matrix only.

In the Bayesian framework, some precision matrix estimation methods can also be applied to covariance matrix with a few adjustments. The underling structure can also be learned by using a conjugate G-inverse Wishart prior that is from the space of positive definite matrices [SG09]. [SG09] use an importance sampling algorithm to approximate the normalizing constant, which is analytically intractable. however, such method can only work for situations with very low dimension. When it comes to the Bayesian penalized methodologies, both the Bayesian graphical lasso by [Wan12] and the SSSL method by [Wan15] can be used for covariance matrix as well, because they do not depend on the underlying undirected or bi-directed graphs.

To serve the demand estimation for MVP models, the Bayesian graphical lasso developed by [Wan12] would be a good choice as long as there are more observations than parameters, because the computation is fast and it also uses Gibbs sampler, which would not add additional computation burden to the MVP model estimation.

3.3 The Bayesian graphical lasso in MVP models

The Bayesian graphical lasso puts independent exponential prior on diagonal entries, Laplace (double exponential) on off-diagonals, and is subject to positive definiteness restriction. It follows the prior choice as [PC08] and uses a scale mixture of normals to represent Laplace prior [Wan12]. It is well know that lasso estimation will over-shrink large coefficients and under-shrink small coefficients. To solve this, [Wan12] extends the the Bayesian graphical lasso to its adaptive version and the SCAD version. The Bayesian adaptive graphical lasso (BAGL) places different shrinkage parameters λ_{ij} on different off-diagonal elements. The distribution of λ_{ij} conditional on Ω is:

$$\lambda_{ij}|\Omega \sim G_A(1+r, |\omega_{ij}|+s)$$

where r and s are hyperparameters. [Wan12] uses $r = 10^{-2}$ and $s = 10^{-6}$, and shows that this choice provides good performance. From the conditional distribution of shrinkage parameters, we can see that the shrinkage parameters can be updated within the MCMC process without cross-validation. The SCAD version, however, requires cross-validation to choose tuning parameter, which is very costly in the Bayesian framework. [Wan12] also shows that the BAGL estimation has very good performance in all of their simulated cases, and therefore I use the BAGL method in this paper. Because the BAGL also uses Gibbs sampler, it requires a small adjustment from the original estimation process. Instead of drawing the precision matrix directly from the Wishart distribution as in the original Gibbs sampler process, I need to divide it into two steps: first update λ_{ij} from $G_A(1+r, |\omega_{ij}|+s)$, and second update precision matrix following the block Gibbs sampler algorithm from [Wan12]. It might seem that it would take a long time to update λ_{ij} for all elements, but it can be easily done in one step by rearranging $\left(\omega_{ij}\right)_{i<j}$ as a vector and drawing $\left(\lambda_{ij}\right)_{i<j}$ all together.

I apply the BAGL method to demand estimation with WCPSS data. Because the high school estimation converges under the normal estimation in Chapter 2, I only estimate the demand for middle and elementary school applicants under heterogeneous sophistical model. The estimation results for elementary and middle school applicants are shown in Tables 3.1 and 3.2 respectively. It is a little hard to compare the results of elementary estimation, as there is no heterogeneous sophistication level estimation for elementary schools in Chapter 2. There are two changes made here, covariance matrix and sophistication levels. Even though I cannot isolate the exact effects from covariance estimation, I can compare the differences among estimation results across school levels. We can see that the error standard deviations increase greatly for both middle and high schools after introducing heterogeneous sophistication levels. However, comparing the results from Tables 3.1 and Table 2.10 in Chapter 2, we can find that most of the error standard deviations do not increase and only a few of them increase slightly after introducing heterogeneous sophistication levels. It suggests the BAGL method leads to more robust estimation for elementary schools. In Tables 2.13 and 3.2, we can find a significant decrease in the error standard deviations. The estimate of teacher turnover rate changes from positive to negative with a smaller standard deviation. The general popularity of schools do not change much, for example Carnage Middle still has the highest average utility, followed by Centennial Campus Middle. Schools with close popularity might have slightly different order under the two approaches. In general, by using the BAGL method within MVP model estimation, I get more robust estimates. Remember that the penalties only apply to the off-diagonal elements, but not the diagonal elements under BAGL. The smaller standard deviations mainly come from better convergence and avoidance of explosive off-diagonal elements. In general, it works well to introduce the BAGL into MVP model estimation without any additional computation burden.

3.4 Bayesian structure learning in graphical models

The Bayesian graphical lasso does not necessarily lead to sparsity and it cannot handle the situation where there are more parameters than number of observations. Several papers have tried to introduce graphical structure to this approach, but they usually come with heavy computational burdens. The posteriors are often intractable, due to the normalizing constants and shrinkage parameters. There are two traditional types of approximation methods. The first type is based on MCMC, such as reverse jump MCMC and birth-death MCMC in trans-dimensional MCMC methodologies. They usually require heavy computation and can only work for low dimension cases, as mentioned previously. The second type of approximation methods are large sample methods, such as Laplace approximation. Large sample methods make a drastic approximation by modeling the posteriors over all parameters as normal and require the computation of Hessian, which makes it computationally expensive [Att00]. With the development of machine learning, recent papers start to use variational Bayes to approximate posterior distribution and ease the computation (e.g. [Att00];

[MM09]; [Che]). Variational Bayes approximates the posterior distribution by choosing hyperparameters that minimize the Kullback-Leibler divergence between the actual posterior and the factorized approximation. [MM09] build a two-step estimation. In the first step, they estimate the graphical structure through a pseudo-likelihood model. In the second step, they estimate parameters of the precision matrix given learned graphical structure. [Che] develop a valid generative model for the data, and their method can learn graphical structure and estimate parameters in one step. Such structure learning methods work very efficiently for high dimensional covariance. However, it uses approximations, but not the exact posterior distributions, which makes it problematic to be used in MVP model estimation. MCMC methods provide guarantees of producing exact samples from the target density, even though the computation is more intense. On the other side, while variational Bayes provides easier and faster computations, it can only find density that's close to the target. Therefore, MCMC methods are more suited for smaller data sets and scenarios where the precision of the sampling outweighs the cost of computation, and variational Bayes is better for the situation where data sets are larger and where computation speed outweighs the accuracy of results [Ble17]. For demand estimation, covariance estimation is only one of many steps, and it could affect the final demand estimation results greatly even with a small estimation bias. The accuracy is therefore very important here. Another factor to consider is that we usually do not have large sample for demand estimation. Therefore, MCMC would be a better option for MVP model demand estimation, especially that I do not have large data sets in my sample. For those who have really large market, where number of parameters is larger than the number of observations, the variational Bayes methods would certainly be a good candidate, particularly the Markov Chain Variational Inference (MCVI) developed by [Sal15], which provides fast posterior approximation with the option of trading off additional computation for additional accuracy, although the application of MCVI in graphical models has not yet been studied.

3.5 Conclusion

In this chapter, I introduce Bayesian (adaptive) graphical lasso to multivariate probit (MVP) models for demand estimation. I first summarize newly developed methods for graphical structure learning and estimation, especially the methodologies suited for high dimensional data in both the classical and Bayesian frameworks. For the specific problem of demand estimation with the MVP model, I find the Bayesian (adaptive) graphical lasso works well. The results show that this method generates reliable estimates, improves convergence, and works well for markets up to a moderate size. I further discuss the limitations of this method and the potential of other methods that could conduct structure learning at the same time. Because I have relatively small sample and the accuracy of the covariance matrix estimation is important, I use the Bayesian (adaptive) graphical lasso. For situations where there are larger samples and computation speed outweighs accuracy, we may consider Variational Bayesian methods.

Table 3.1 Estimation of elementary schools, Bayesian graphical lasso

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Brentwood E	-1.76	0.67	1.53	-2.08	2.88	3.27	1.66	0.75	0.22	1.08	10.14
SD	2.57	0.60	2.65	0.62	1.39	1.03	1.29	1.10	1.40	1.39	4.64
Brooks E	8.91	-3.19	-2.43	-0.86	0.51	-1.44	0.00	-2.04	-5.08	2.00	6.09
SD	3.38	0.83	2.79	0.72	1.28	0.80	1.28	1.20	1.76	1.35	2.35
Bugg E	0.96	-0.96	-2.21	0.15	2.80	5.16	0.60	1.31	1.54	0.66	7.84
SD	2.46	0.69	1.83	0.53	0.85	1.49	1.36	0.96	1.16	1.39	3.67
Combs E	7.10	-1.42	2.30	-1.17	1.27	1.49	2.88	0.42	-4.70	2.63	8.29
SD	2.52	0.91	2.22	0.64	0.95	0.91	1.43	1.11	1.40	1.51	4.13
Conn E	0.13	0.40	1.85	-0.99	0.97	5.00	2.46	0.56	-0.95	1.82	9.41
SD	2.20	0.70	2.18	0.63	1.16	1.11	1.46	1.13	1.39	1.20	3.38
Douglas E	7.40	-7.45	1.94	-0.30	-2.73	-0.35	-1.17	-2.70	-2.56	1.77	6.94
SD	3.30	1.06	2.85	0.84	1.60	1.14	1.88	1.64	1.89	1.69	2.78
Farmington E	4.80	-5.72	-4.47	-0.81	2.11	-1.08	-0.66	-0.34	-1.86	-5.02	4.75
SD	3.80	0.87	1.51	0.73	0.91	1.08	1.99	1.53	1.32	2.64	2.61
Fox Road E	-4.12	-1.56	-6.09	-1.83	6.84	5.03	3.11	3.72	-6.91	3.01	8.13
SD	4.28	0.94	4.50	0.90	1.40	1.57	1.53	2.22	2.09	1.62	2.84
Fuller E	-0.25	1.13	6.59	-1.12	17.34	2.95	5.88	0.33	-1.88	-1.33	10.36
SD	2.47	0.84	2.17	1.13	1.09	1.21	1.72	1.72	1.73	2.22	3.00
Green E	-7.61	0.44	3.13	-1.74	-1.77	2.64	-2.44	-0.30	-1.17	1.71	14.26
SD	4.23	0.99	3.49	1.11	2.31	1.21	1.91	1.70	2.29	1.84	4.65
Hunter E	-3.30	0.78	5.65	0.15	1.83	3.67	3.74	0.05	-0.41	-1.54	11.04
SD	2.33	0.91	2.16	0.76	1.11	1.03	1.53	1.26	1.66	1.53	3.68
Joyner E	4.72	-3.82	1.68	-1.26	-2.58	1.63	2.97	3.14	-2.20	2.05	9.33
SD	2.66	0.91	3.85	0.85	1.63	0.96	1.59	1.41	1.78	1.59	3.49
Kingswood E	-2.48	-5.55	4.86	0.76	-7.75	0.48	-2.60	3.89	-1.77	7.71	13.79
SD	3.33	1.33	3.21	1.22	1.60	1.98	2.84	1.80	2.15	2.13	4.96

Tables 3.1 (continued)

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Millbrook E	-4.03	-1.17	-4.34	-1.19	3.22	4.34	2.26	1.97	-0.58	1.98	10.46
SD	2.61	0.67	2.28	0.75	1.45	1.30	1.50	1.25	1.42	1.38	4.11
Partnership E	-9.46	-0.14	-7.08	-1.29	-3.17	4.75	3.44	-0.29	-2.02	3.66	12.40
SD	6.86	0.93	8.61	0.85	2.05	1.38	1.69	1.56	2.03	1.68	5.08
Poe E	-7.12	-1.10	2.96	-1.32	4.20	6.98	3.25	3.29	-0.19	1.89	13.35
SD	2.71	0.91	3.16	0.89	1.58	1.42	2.05	1.68	1.77	1.92	3.69
Powell E	-6.44	-1.42	2.28	-0.71	-0.48	5.07	2.45	1.08	3.31	-1.10	12.16
SD	3.34	1.03	2.81	0.91	1.54	1.52	2.05	1.46	1.62	1.91	5.18
Smith E	-9.09	-1.81	-4.92	-0.80	-12.52	3.05	-7.48	1.85	-3.81	0.29	12.56
SD	5.02	1.39	10.58	1.42	10.47	2.18	6.74	2.37	3.84	4.31	4.97
Underwood E	-1.39	-0.35	1.58	-0.60	5.96	-1.30	1.59	-4.77	-8.12	0.72	7.93
SD	3.90	0.89	3.35	0.98	1.13	1.13	1.72	2.53	3.27	1.81	3.64
Washington E	6.38	-2.27	3.04	-2.18	13.35	1.10	4.02	-0.39	-1.31	1.95	7.40
SD	2.70	0.79	1.94	0.91	0.94	0.96	1.34	1.32	1.45	1.53	2.31
Wendell E	-0.81	0.67	-9.08	0.96	-0.53	2.94	3.30	-4.44	-1.87	-1.05	10.50
SD	2.60	1.08	3.07	0.76	1.57	0.98	1.70	1.69	3.44	1.73	4.02
Wiley E	8.46	-2.90	2.18	-0.64	1.50	0.06	3.10	-1.41	-0.99	1.67	5.65
SD	3.91	0.74	1.91	0.63	1.08	0.75	1.18	1.15	1.23	1.28	2.39
Zebulon E	-2.40	-1.05	-1.75	-1.05	4.52	6.72	2.94	-12.25	2.10	3.45	10.98
SD	2.46	1.09	5.93	1.35	3.50	2.20	2.29	3.68	3.71	2.25	3.71

Panel B School-constant estimation results

	TradCal	Turnover	SPG
Estimates	-1.97	-0.02	0.17
SD	7.08	0.44	0.11

Notes: See Tables 2.7 and 2.10 for definitions.

Table 3.2 Estimation of middle schools, Bayesian graphical lasso

Panel A School-specific estimation results

School	Intercept	NGE	AG	Female	Asian	Black	Other	Hisp	LEP	SPED	Unobs
Carnage M	-3.40	0.87	-0.05	-2.28	14.50	2.32	2.18	1.06	1.96	-1.91	67.13
SD	7.31	6.84	2.77	2.17	2.60	2.26	2.70	2.64	2.47	2.33	45.18
Carroll M	8.42	-7.08	-5.64	0.30	-6.37	4.15	-0.28	-1.44	-0.02	1.45	147.48
SD	5.78	17.39	7.18	5.28	6.08	5.58	6.84	6.68	6.27	5.90	146.99
Centennial M	23.64	-10.39	-22.57	-2.67	-10.53	-2.45	3.80	-4.03	-3.79	3.52	84.06
SD	6.18	4.70	3.78	2.71	3.91	3.20	5.03	4.86	4.55	3.95	45.85
E Garner M	-0.95	10.80	5.00	-0.49	2.33	-2.28	4.12	4.63	5.13	-1.89	77.44
SD	6.79	3.42	2.63	2.10	2.82	2.62	3.81	3.97	3.33	3.47	40.47
E Millbrook M	-6.91	-7.65	-3.37	1.37	4.93	8.18	0.08	4.38	-0.86	5.11	373.86
SD	6.45	7.07	5.98	5.89	6.45	6.20	6.73	6.69	6.54	6.66	227.94
Ligon M	-0.99	5.92	6.40	-4.53	-8.37	-11.23	-7.20	-2.75	-0.84	-5.64	362.56
SD	5.84	7.22	6.20	6.06	6.88	6.59	7.11	7.23	6.92	6.96	213.29
Martin M	11.12	0.03	-10.08	-4.48	2.47	-25.43	-8.24	-12.35	2.02	-1.30	121.53
SD	6.18	6.11	5.16	4.32	5.62	5.85	7.01	6.75	6.39	6.17	70.72
Moore Square M	-16.25	1.56	-29.17	-3.73	-8.55	13.04	-0.85	7.16	-2.31	9.63	117.41
SD	6.94	6.90	5.37	4.62	6.23	5.41	7.50	6.99	6.52	6.13	58.32
Zebulon M	-18.62	-19.60	-9.22	4.55	5.90	8.75	7.65	-0.29	2.01	-2.42	238.55
SD	5.15	4.59	4.05	3.90	4.52	4.22	4.85	5.09	4.87	4.45	142.19

Panel B School-constant estimation results

	TradCal	Turnover	SPG
Estimates	-11.28	-1.15	0.53
SD	7.00	0.54	0.19

Notes: Tables 2.7 and 2.10 for definitions.

BIBLIOGRAPHY

[Abd15] Abdulkadiroglu, A. et al. *The Welfare Effects of Coordinated Assignment: Evidence from the NYC HS Match*. Working Paper 21046. National Bureau of Economic Research, 2015.

[Abd17] Abdulkadiroglu, A. et al. *Minimizing Justified Envy in School Choice: The Design of New Orleans' OneApp*. Tech. rep. National Bureau of Economic Research, 2017.

[Abd09] Abdulkadiroğlu, A. et al. "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match". *American Economic Review* **99** (2009), pp. 1954–1978.

[Abd11] Abdulkadiroğlu, A. et al. "Resolving Conflicting Preferences in School Choice: the Boston Mechanism Reconsidered". *American Economic Review* **101**.1 (2011), pp. 399–410.

[AS16] Agarwal, N. & Somaini, P. *Demand Analysis using Strategic Reports: An application to a school choice mechanism*. Tech. rep. National Bureau of Economic Research, 2016.

[Aja13] Ajayi, K. F. "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana". mimeo. 2013.

[AC93] Albert, J. H. & Chib, S. "Bayesian Analysis of Binary and Polychotomous Response Data". *Journal of the American Statistical Association* **88**.422 (1993), pp. 669–679.

[AR99] Allenby, G. M. & Rossi, P. "Marketing models of consumer heterogeneity". *Journal of Econometrics* **89** (1999), pp. 57–78.

[AN17] Ashlagi, I. & Nikzad, A. "What matters in school choice tie-breaking? How competition guides design". mimeo. 2017.

[Att00] Attias, H. "A Variational Bayesian Framework for Graphical Models". *In Advances in Neural Information Processing*. MIT Press, 2000.

[BG14] Banerjee, S. & Ghosal, S. "Posterior convergence rates for estimating large precision matrices using graphical models". *Electronic Journal of Statistics* **8**.2 (2014), pp. 2111–2137.

[BG15] Banerjee, S. & Ghosal, S. "Bayesian structure learning in graphical models". *Journal of Multivariate Analysis* **136** (2015), pp. 147–162.

[BH10] Berry, S. T. & Haile, P. A. "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers". mimeo. 2010.

[BH14] Berry, S. T. & Haile, P. A. "Identification in differentiated products markets using market level data". *Econometrica* (2014).

[BL08] Bickel, P. J. & Levina, E. "Covariance regularization by thresholding". *The Annals of Statistics* **36**.6 (2008), pp. 2577–2604.

[BT11] Bien, J. & Tibshirani, R. "Sparse estimation of a covariance matrix". *Biometrika* **98**.4 (2011), pp. 807–820.

[Ble17] Blei, D. M. et al. "Variational Inference: A Review for Statisticians". *Journal of the American Statistical Association* **112**.518 (2017), pp. 859–877.

[Bol97] Bolduc, D. et al. "Multinomial Probit Estimation of Spatially Interdependent Choices: An Empirical Comparison of Two New Techniques". *International Regional Science Review* **20** (1997), pp. 77–101.

[CL11] Cai, T. & Liu, W. "Adaptive Thresholding for Sparse Covariance Matrix Estimation". *Journal of the American Statistical Association* **106**.494 (2011), pp. 672–684.

[Cal10] Calsamiglia, C. et al. "Constrained School Choice: An Experimental Study". *American Economic Review* **100**.4 (2010), pp. 1860–1874.

[Cal16] Calsamiglia, C. et al. "Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives". mimeo. 2016.

[CS09] Carvalho, C. M. & Scott, J. G. "Objective Bayesian model selection in Gaussian graphical models". *Biometrika* **96** (2009), pp. 1–16.

[Cha07] Chaudhuri, S. et al. "Estimation of a Covariance Matrix with Zeros". *Biometrika* **94**.1 (2007), pp. 199–216.

[Che] Chen, M. et al. "Bayesian Learning of Sparse Gaussian Graphical Models".

[CS06] Chen, Y. & Sönmez, T. "School choice: an experimental study". *Journal of Economic Theory* **127**.1 (2006), pp. 202–231.

[CW93] Cox, D. R. & Wermuth, N. "Linear dependence represented by chain graphs". *Statistical Science* **8**.3 (1993), pp. 204–218.

[Dem72] Dempster, A. P. "Covariance selection". *Biometrics* **28**.1 (1972), pp. 157–175.

[Dob11] Dobra, A. et al. "Bayesian inference for general Gaussian graphical models with application to multivariate lattice data". *Journal of the American Statistical Association* **106**.496 (2011), pp. 1418–1433.

[Dur17a] Dur, U. et al. "Identifying the Harm of Manipulable School-Choice Mechanisms". *American Economic Journal: Economic Policy* (2017).

[Dur17b] Dur, U. et al. "Sequential School Choice: Theory and Evidence from the Field and Lab". mimeo. 2017.

[ES06] Ergin, H. & Sönmez, T. “Games of School Choice under the Boston Mechanism”. *Journal of Public Economics* **90** (2006), pp. 215–237.

[Fac15] Fack, G. et al. “Beyond Truth-Telling: Preference Estimation with Centralized School Choice”. mimeo. 2015.

[Fan09] Fan, J. et al. “Network exploration via the adaptive LASSO and SCAD penalties”. *The Annals of Applied Statistics* **3**.2 (2009), pp. 521–541.

[Fri08] Friedman, J. et al. “Sparse inverse covariance estimation with the graphical lasso”. *Biostatistics* **9**.3 (2008), pp. 432–441.

[Gew91] Geweke, J. “Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities”. 1991.

[GG99] Giudici, P. & Green, P. J. “Decomposable graphical Gaussian model determination”. *Biometrika* **86**.4 (1999), pp. 785–801.

[Gre95] Green, P. J. “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination”. *Biometrika* **82**.4 (1995), pp. 711–732.

[Guo11] Guo, J. et al. “Joint estimation of multiple graphical models”. *Biometrika* (2011).

[Haa16] Haan, M. de et al. “The performance of school assignment mechanisms in practice”. mimeo. 2016.

[Haj90] Hajivassiliou, V. A. “Smooth Estimation Simulatio of Panel Data LDV Models”. 1990.

[He16] He, Y. “Gaming the Boston School Choice Mechanism in Beijing”. mimeo. 2016.

[Hwa16] Hwang, S. I. M. “A Robust Redesign of High School Match”. mimeo. 2016.

[Jon05] Jones, B. et al. “Experiments in Stochastic Computation for High-Dimensional Graphical Models”. *Statistical Science* **20**.4 (2005), pp. 388–400.

[Kap17] Kapor, A. et al. “Heterogeneous Beliefs and School Choice Mechanisms”. mimeo. 2017.

[Kau96] Kauermann, G. “On a Dualization of Graphical Gaussian Models”. *Scandinavian Journal of Statistics* **23**.1 (1996), pp. 105–116.

[Kea94] Keane, M. P. “A Computationally Practical Simulation Estimator for Panel Data”. *Econometrica* **62**.1 (1994), pp. 95–116.

[Lau96] Lauritzen, S. L. *Graphical models*. Clarendon Press, 1996.

[Len13] Lenkoski, A. “A Direct Sampler for G-Wishart Variates”. *Stat* **2**.119-128 (2013).

[Lew00] Lewbel, A. “Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables”. *Journal of Econometrics* **97**.1 (2000), pp. 145–177.

[MM09] Marlin, B. M. & Murphy, K. P. “Sparse Gaussian graphical models with unknown block structure”. *Proceedings of ICML ’09 Proceedings of the 26th Annual International Conference on Machine Learning* (2009).

[MR94] McCulloch, R. & Rossi, P. E. “An exact likelihood analysis of the multinomial probit model”. *Journal of Econometrics* **64** (1994), pp. 207–240.

[McF89] McFadden, D. “A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration”. *Econometrica* **57**.5 (1989), pp. 995–1026.

[MB06] Meinshausen, N. & Bühlmann, P. “High-Dimensional Graphs and Variable Selection with the Lasso”. *The Annals of Statistics* **34**.3 (2006), pp. 1436–1462.

[Mir08] Miralles, A. “School Choice: The Case for the Boston Mechanism”. Boston University, unpublished mimeo. 2008.

[MW15] Mohammadi, A. & Wit, E. C. “Bayesian Structure Learning in Sparse Gaussian Graphical Models”. *Bayesian Analysis* **10**.1 (2015), pp. 109–138.

[Nar16] Narita, Y. “Match or Mismatch: Learning and Inertia in School Choice”. mimeo. 2016.

[PP08] Pais, J. & Pintér, Ágnes. “School choice and information: An experimental study on matching mechanisms”. *Games and Economic Behavior* **64**.1 (2008), pp. 303–328.

[PC08] Park, T. & Casella, G. “The Bayesian Lasso”. *Journal of the American Statistical Association* **103**.482 (2008), pp. 681–686.

[Pat16] Pathak, P. A. “What really matters in designing school choice mechanisms”. *Advances in Economics and Econometrics, 11th World Congress of the Econometric Society*. 2016.

[PS08] Pathak, P. A. & Sönmez, T. “Leveling the playing field: Sincere and sophisticated players in the Boston mechanism”. *American Economic Review* **98**.4 (2008), pp. 1636–1652.

[Ros96] Rossi, P. E. et al. “The Value of Purchase History Data in Target Marketing”. *Marketing Science* **15**.4 (1996), pp. 321–340.

[Rot08] Rothman, A. J. et al. “Sparse permutation invariant covariance estimation”. *Electronic Journal of Statistics* **2** (2008), pp. 494–515.

[Rot09] Rothman, A. J. et al. “Generalized Thresholding of Large Covariance Matrices”. *Journal of the American Statistical Association* **104**.485 (2009), pp. 177–186.

[Rov00] Roverato, A. “Cholesky decomposition of a hyper inverse Wishart matrix”. *Biometrika* **87**.1 (2000), pp. 99–112.

[Rov02] Roverato, A. “Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models”. *Scandinavian Journal of Statistics* **29**.3 (2002), pp. 391–411.

[Sal15] Salimans, T. et al. “Markov Chain Monte Carlo and Variational Inference: Bridging the Gap”. *Proceedings of the 32nd International Conference on Machine Learning* (2015).

[SG09] Silva, R. & Ghahramani, Z. “The Hidden Life of Latent Variables: Bayesian Learning with Mixed Graph Models”. *Journal of Machine Learning Research* **10** (2009), pp. 1187–1238.

[Tra09] Train, K. E. “Discrete Choice Methods with Simulation”. Cambridge University Press, 2009. Chap. Bayesian Procedures, pp. 282–314.

[Tro12] Troyan, P. “Comparing school choice mechanisms by interim and ex-ante welfare”. *Games and Economic Behavior* 75 (2012), pp. 936–947.

[Wan12] Wang, H. “Bayesian Graphical Lasso Models and Efficient Posterior Computation”. *Bayesian Analysis* **7**.4 (2012), pp. 867–886.

[Wan15] Wang, H. “Scaling It Up: Stochastic Search Structure Learning in Graphical Models”. *Bayesian Analysis* **10**.2 (2015), pp. 351–377.

[WC10] Wang, H. & Carvalho, C. M. “Simulation of hyper-inverse Wishart distributions for non-decomposable graphs”. *Electronic Journal of Statistics* **4** (2010), pp. 1470–1475.

[WL12] Wang, H. & Li, S. Z. “Efficient Gaussian graphical model determination under G-Wishart prior distributions”. *Electronic Journal of Statistics* **6** (2012).

[WCP16] WCPSS. “Wake County Public School System District Facts 2016-2017”. `http://www.wcpss.net/domain/100`. 2016.

[Wer06] Wermuth, N. et al. “Covariance chains”. *Bernoulli* **12**.5 (2006), pp. 841–862.

[Wit11] Witten, D. M. et al. “New Insights and Faster Computations for the Graphical Lasso”. *Journal of Computational and Graphical Statistics* **20**.4 (2011), pp. 892–900.

[Yua08] Yuan, M. “Efficient Computation of $l1$ Regularized Estimates in Gaussian Graphical Models”. *Journal of Computational and Graphical Statistics* **17**.4 (2008), pp. 809–826.

[YL07] Yuan, M. & Lin, Y. “Model selection and estimation in the Gaussian graphical model”. *Biometrika* **94**.1 (2007), pp. 19–35.

APPENDICES

APPENDIX

A

ESTIMATION DETAILS OF GIBBS SAMPLER

A.1 Homogeneous Sophistication

A.1.1 Priors and Initial Values

To conduct Bayesian estimation, we need to set initial values of utilities for each student so that they follow the rule of maximizing expected utility. This could be easily done by three steps: first setting utility for one school, second solving the inequalities of maximizing expected utility with the given utility to get the bounds of utilities for other schools, and third drawing the random numbers within these bounds. Denote initial values as U_0 and $U_0 = (u_1^0, \ldots, u_N^0)$. The priors for θ (all coefficients) and Σ (variance-covariance matrix of the error term) are set as following:

$$\theta \sim \mathrm{N}(\bar{\theta}, A^{-1}),$$

$$\Sigma \sim \mathrm{IW}(\upsilon, V)$$

where θ=β when only characteristics with schools varying estimates are used and θ=[β δ] when school level characteristics with schools constant estimates are included; and IW is the inverse

Wishart distribution. Let $G = \Sigma^{-1}$. We could alternatively set prior of G with Wishart distribution. To make sure the influence from the prior would not dominate the influence from samples, we use diffuse priors:

$$\theta: \quad \bar{\theta} = 0,\ A^{-1} = 100 \times I,$$

$$\Sigma: \quad \upsilon = 50,\ V = I.$$

We tried two different sets of initial values and some more diffuse priors. There was not much difference with estimation results in either case.

A.1.2 Steps of Gibbs Sampler

Let IX_i denotes all the independent variables included in the model for student i: $\mathrm{IX}_i = X_i$ for the baseline model and $\mathrm{IX}_i = [X_i,\ z]$ when school level characteristics with schools constant estimates are included, where X_i and z are constructed as in Section 4. We use the following model to represent both equation (8) and equation (9) in Section 4:

$$u_i = \mathrm{IX}_i\theta - D_i + \varepsilon_i \tag{A.1}$$

Define $w_i = u_i + D_i$. Then the estimation equation A.1 is equivalent to:

$$w_i = \mathrm{IX}_i\theta + \varepsilon_i \tag{A.2}$$

We also need to construct a big matrix IX as the way to construct X ($SN \times SJ$ matrix) and and Z ($SN \times Q$ matrix) as in Section 4, where J is number of individual characteristics and Q is the number of school characteristics. They are constructed as below:

$$\mathrm{IX} = \begin{bmatrix} \mathrm{IX}_1 \\ \mathrm{IX}_2 \\ \vdots \\ \mathrm{IX}_N \end{bmatrix},\ X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix},\ \text{and } Z = \begin{bmatrix} z \\ z \\ \vdots \\ z \end{bmatrix}.$$

After setting up the matrix for estimation, the rest Bayesian process follows strictly as in [MR94], where the nth loop of the Gibbs Sampler is as below:

1. Draw $\theta^n | U^{n-1}, \Sigma^{n-1}$from a $\mathrm{N}(\hat{\theta},\ \Sigma_\theta)$, where$\hat{\theta}$ and $\Sigma_\theta$are defined in the following system:

   Pre-multiply both sides of equation A.2 by C^T, where C is the Cholesky root of G and $G = CC^T$. This will transform the original system to a system with $\mathrm{N}(0,\ I)$ errors:

$$C^T w_i = C^T \mathrm{IX}_i\theta + C^T\varepsilon_i. \tag{A.3}$$

Rewrite equation A.3 as:

$$w_i^* = \text{IX}_i^* \theta + \varepsilon_i^*, \text{ and } \varepsilon \sim \text{N}(0, I_S) \tag{A.4}$$

or in stacked form,

$$W^* = \text{IX}^* \theta + \varepsilon^*, \text{ and } \varepsilon \sim \text{N}(0, I_{NS}) \tag{A.5}$$

Keep in mind that W^* needs to be formed using U^{n-1}. Then the conditional posterior distribution of θ is $N(\hat{\theta}, \Sigma_\theta)$, where

$$\Sigma_\theta = (\text{IX}^{*^T} \text{IX}^* + A)^{-1}, \qquad \hat{\theta} = \Sigma_\theta(\text{IX}^{*^T} W^* + A\bar{\theta}).$$

2. Draw $\Sigma^n | U^{n-1}, \theta^n$ from a inverse Wishart distribution $\text{IW}(\upsilon + N, V + \sum_{i=1}^{N} \varepsilon_i \varepsilon_i^T)$, where

$$\varepsilon_i = w_i^{n-1} - \text{IX}_i \theta^n = u_i^{n-1} + D_i - \text{IX}_i \theta^n.$$

3. Draw $U^n | \theta^n, \Sigma^n, R$, where R denotes the rankings submitted by students. The utilities need to be updated by iterating over students and schools. For a student i who submitted a ranking R_i, we need to update her utilities in all schools one by one given the utilities of the rest schools. Her utility for school s given utilities of the other schools, $u_{is}^n | u_{i1}^n, \ldots, u_{i(s-1)}^n, u_{i(s+1)}^{n-1}, \ldots, u_{iS}^{n-1}$, would be drawn from a truncated normal

$$\text{TN}(m_{is}, \tau_{is}^2, \text{lower}_{is}, \text{upper}_{is})$$

where lower_{is} is the lower bound of u_{is}, upper_{is} is the upper bound of u_{is}, and m_{is}, τ_{is}^2 are as below:

$$G = \begin{bmatrix} \Sigma_{(-s)(-s)} & \sigma_{(-s)s} \\ \sigma_{s(-s)} & \sigma_{ss} \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma_{(-s)(-s)}^{-1} + FE^{-1}F^T & -FE^{-1} \\ -E^{-1}F^T & E^{-1} \end{bmatrix} \tag{A.6}$$

where the first bracket of equation A.6 is a convenient form of G permuted by putting the rows and columns of G so that the sth row and column are the last. E, F, m_{is}, and τ_{is}^2 are as below:

$$E = \sigma_{ss} - \sigma_{s(-s)} \Sigma^{-1}_{(-s)(-s)} \sigma_{(-s)s},$$
$$F = \Sigma^{-1}_{(-s)(-s)} \sigma_{(-s)s},$$
$$\tau^2_{is} = E,$$
$$m_{is} = \text{ix}^T_{is}\theta - d_{is} + F^T(u_{i(-s)} + D_{i(-s)} - \text{IX}_{i(-s)}\theta)$$

where ix_{is} is the sth row of IX_i, d_{is} is the sth row of D_i, $\text{IX}_{i(-s)}$ is the sub-matrix created by deleting the sth row from the IX_i, $\text{IX}_{i(-s)}$ is the sub-vector created by deleting the sth element from the D_i, and $u_{i(-s)} = (u^n_{i1}, \ldots, u^n_{i(s-1)}, u^{n-1}_{i(s+1)}, \ldots, u^{n-1}_{iS})$. In particular,

$$\text{ix}^T_{is}\theta = \sum_{j=1}^{J} \beta_{sj} x_{isj} + \sum_{q=1}^{Q} \delta_q z_{sq}.$$

Now, we need to get the upper bound and lower bound of u_{is}. First, set lower_{is} to be $-\infty$ if school s is not listed in the R_i, and 0.000001 if it's listed. upper_{is} is set to be ∞. Then they might be updated by solving the inequalities of maximizing expected utility. Let $\boldsymbol{P}_i = (P_1, \ldots, P_{|\boldsymbol{P}_i|})$ denote the set of all possible non-repeated assignment probabilities for student i with all feasible rankings, where $|\boldsymbol{P}_i|$ is the length of $\boldsymbol{P}_i$ and each element of $\boldsymbol{P}_i$ is a $S \times 1$ vector of assignment probabilities. Let P_{R_i} be the assignment probability of R_i. Then we could get the bounds of utilities of student i in all the schools by solving a system of inequalities:

$$\begin{bmatrix} P^T_{R_i} - P^T_1 \\ P^T_{R_i} - P^T_2 \\ \cdots \\ P^T_{R_i} - P^T_{|\boldsymbol{P}_i|} \end{bmatrix}_{|\boldsymbol{P}_i| \times S} \cdot \begin{bmatrix} u_{i1} \\ u_{i2} \\ \cdots \\ u_{iS} \end{bmatrix} \geqslant \mathbf{0} \tag{A.7}$$

or rewrite it as:

$$\Gamma_i u_i \geqslant 0 \tag{A.8}$$

where Γ_i is the$|\boldsymbol{P}_i| \times S$ matrix in the left bracket on the left-hand side of inequality A.7. To solve the bounds for school s, rewrite the system A.8 as:

$$\Gamma_i(:, 1)u_{i1} + \Gamma_i(:, 2)u_{i3} + \ldots + \Gamma_i(:, S)u_{is} \geqslant 0 \tag{A.9}$$

where$\Gamma_i(:, s)$ means the sth column of Γ_i. Solving for u_{is}:

$$\Gamma_i(:, s)u_{is} \geqslant -\Gamma_{-i}(:, -s)u_{i(-s)} \tag{A.10}$$

where $\Gamma_i(:,-s)$ is the sub-matrix created by removing the sth column of Γ_i. From inequality A.10 we could solve for upper and lower bound:

$$\text{lower}_{is} = \max_{l\in\{l:\Gamma_i(l,s)>0\}} \frac{-\Gamma_{-i}(:,-s)u_{i(-s)}}{\Gamma_i(l,s)},$$

$$\text{upper}_{is} = \min_{l\in\{l:\Gamma_i(:,s)<0\}} \frac{-\Gamma_{-i}(:,-s)u_{i(-s)}}{\Gamma_i(l,s)},$$

where $\Gamma_i(l,s)$ is the lth element of $\Gamma_i(:,s)$. Again, remember that

$$u_{i(-s)} = (u_{i1}^n, \ldots, u_{i(s-1)}^n, u_{i(s+1)}^{n-1}, \ldots, u_{iS}^{n-1}).$$

After updating utilities of all the schools for all the students, all the variables are updated and the nth loop ends.

A.2 Heterogeneous Sophistication

A.2.1 Priors and Initial Values

The priors for a heterogeneous sophistication model follow the sane setup as in the homogeneous sophistication model. The initial values of utility are different, because there are two types of agents and they follow different strategies. We set initial values of utilities for each student such that the utilities of sincere agents follow the equality of truth-telling and utilities of sophisticated agents follow the inequality of maximizing expected utility in Section 3.

A.2.2 Steps of Gibbs Sampler

Most steps of Gibbs Sampler here are the same as the homogeneous sophistication model. The only difference happens to the third step, updating utilities for each student in all schools. There are two types here. For sophisticated agents who maximize expected utility, the update process is the same as in homogeneous sophistication model above. For sincere agents who follow the truth-telling strategy, the update process is different. Utilities are still updated by iterating over students and schools. And the utility of student i in school s given her utilities in all other schools, $u_{is}^n | u_{i1}^n, \ldots, u_{i(s-1)}^n, u_{i(s+1)}^{n-1}, \ldots, u_{iS}^{n-1}$, would still be drawn from a truncated normal $\text{TN}(m_{is}, \tau_{is}^2, \text{lower}_{is}, \text{upper}_{is})$. m_{is} and τ_{is}^2 are also calculated in the same way as in the homogeneous sophistication model. However, the lower bound lower_{is} and the upper bound upper_{is} are updated differently. First, denote U_i^{des} as the vector of utilities of student i that's sorted in a descending order. If student i ranks school s first, then $\text{upper}_{is} = \inf$ and $\text{lower}_{is} = \max\{0, U_{i,-1}^{\text{des}}\}$, where $U_{i,-1}^{\text{des}}$ is the vector after removing the first element from U_i^{des}. If student i ranks school s second, then $upper_{is} = u_{i,1}^{\text{des}}$ and $\text{lower}_{is} = \max\{0, U_{i,-2}^{\text{des}}\}$, where $u_{i,1}^{\text{des}}$ is the first element of U_i^{des} and $U_{i,-2}^{\text{des}}$ is the vector after removing the first two elements from U_i^{des}. If student i ranks school s third, then $upper_{is} = u_{i,2}^{\text{des}}$ and $\text{lower}_{is} = \max\{0, U_{i,-3}^{\text{des}}\}$, where $u_{i,2}^{\text{des}}$ is the second element of U_i^{des} and $U_{i,-3}^{\text{des}}$ is the vector after removing the first three elements from U_i^{des}.

APPENDIX

B

ESTIMATION PLOTS

B.1 Homogeneous Sophistication

Because I run separate estimations for schools at different levels for both with and without school characteristics. There are many plots generated. Here, I would only provide the plots for estimation for high schools with school characteristics, because there are fewest candidates high schools.

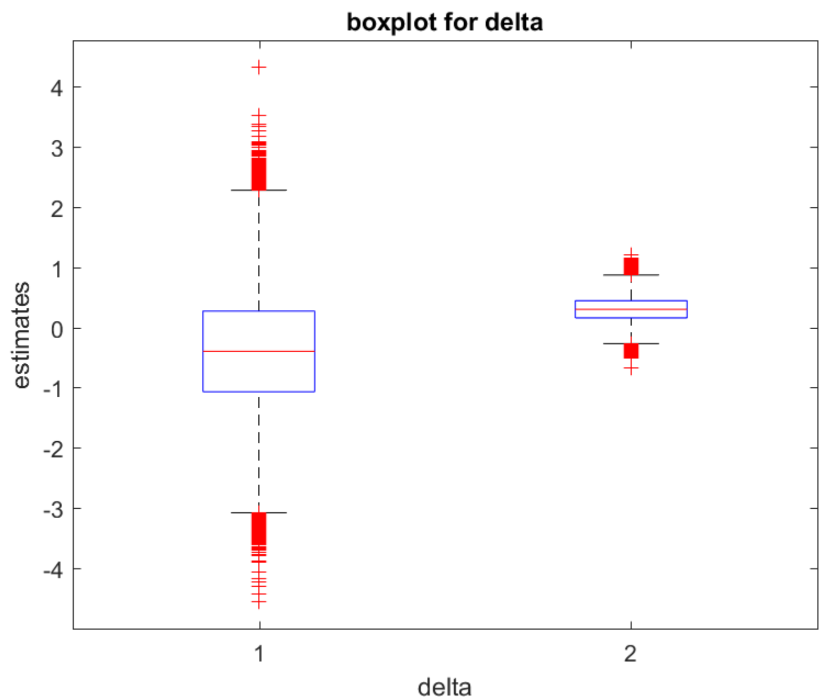

Figure B.1 Boxplot of coefficients for school characteristics, homogeneous sophistication

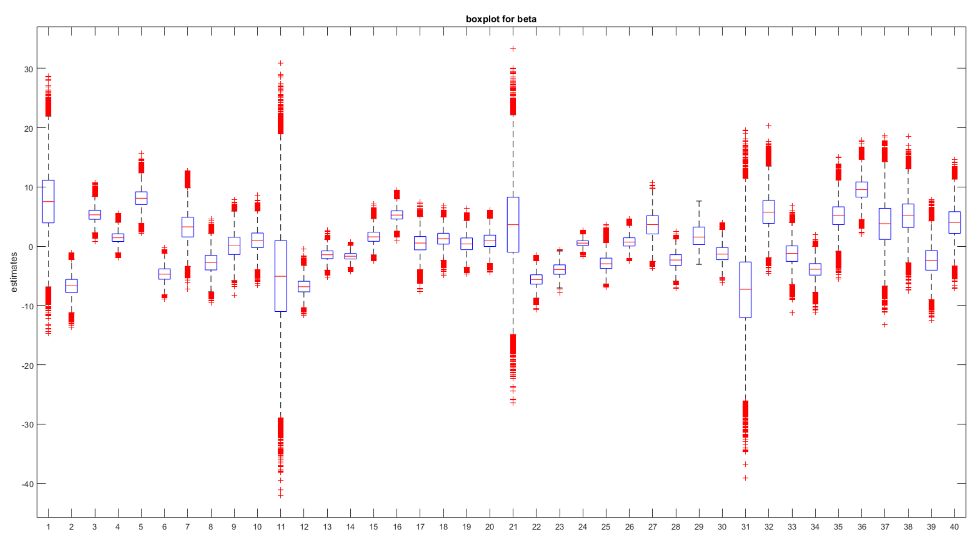

Figure B.2 Boxplot of coefficients for individual characteristics, homogeneous sophistication

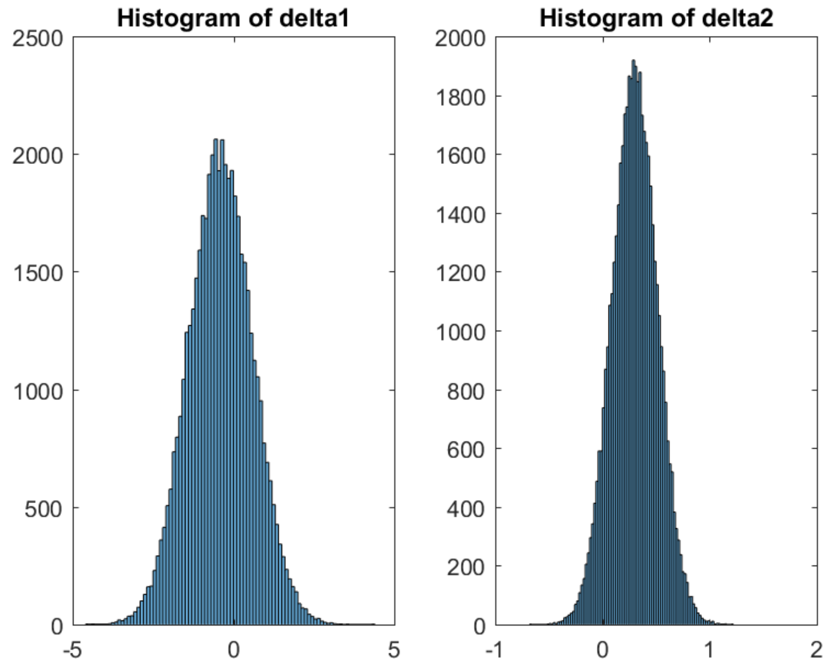

Figure B.3 Histogram of coefficients for school characteristics, homogeneous sophistication

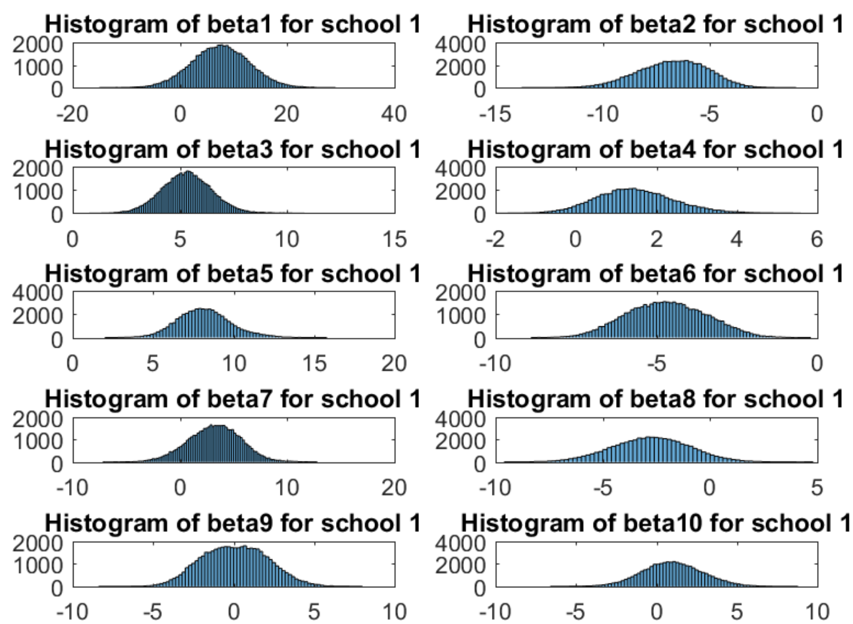

Figure B.4 Histogram of coefficients for individual characteristics in high school 1, homogeneous sophistication

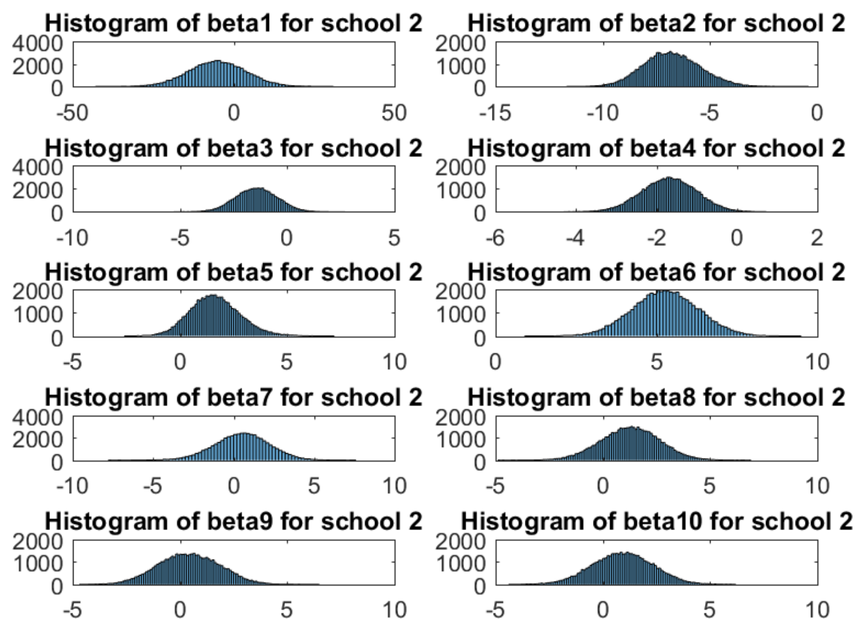

Figure B.5 Histogram of coefficients for individual characteristics in high school 2, homogeneous sophistication

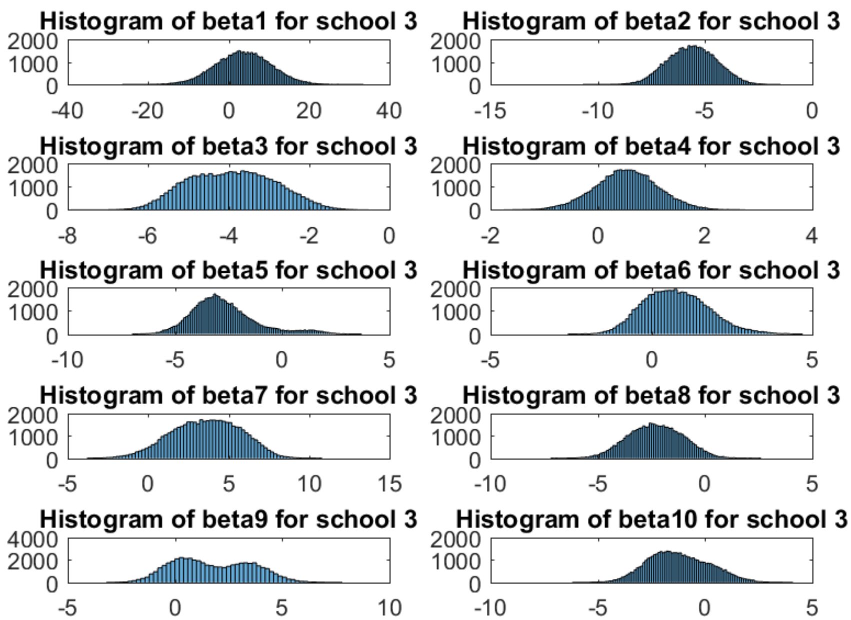

Figure B.6 Histogram of coefficients for individual characteristics in high school 3, homogeneous sophistication

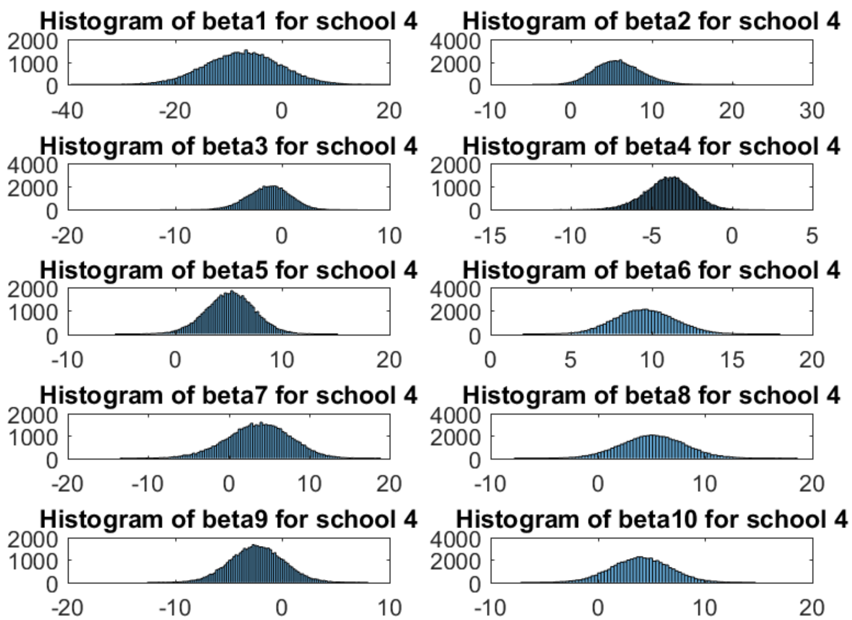

Figure B.7 Histogram of coefficients for individual characteristics in high school 4, homogeneous sophistication

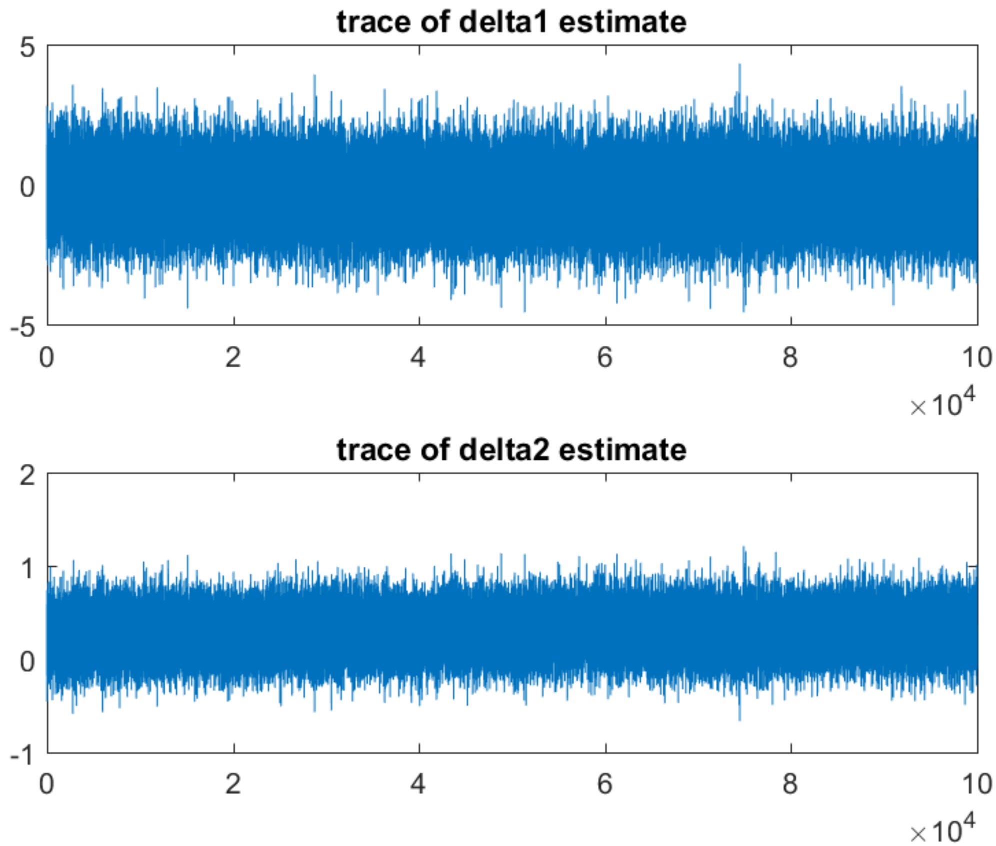

Figure B.8 Trace of coefficients for school characteristics, homogeneous sophistication

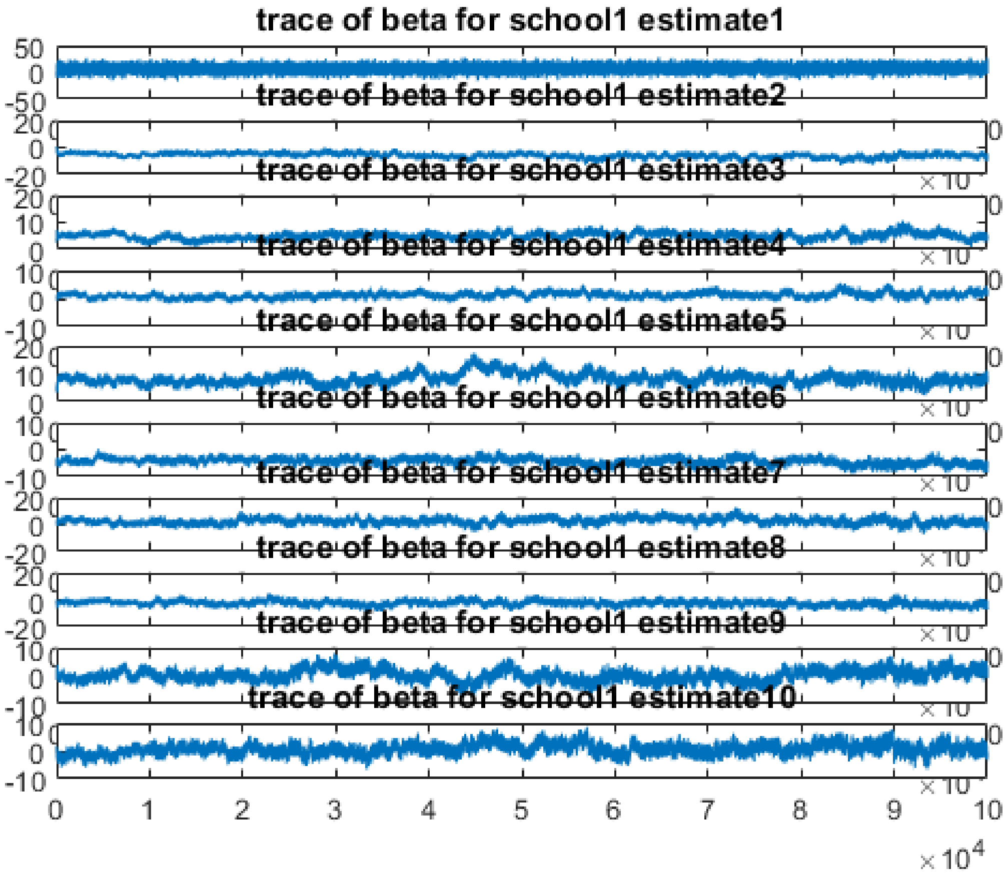

Figure B.9 Trace of coefficients for individual characteristics in high school 1, homogeneous sophistication

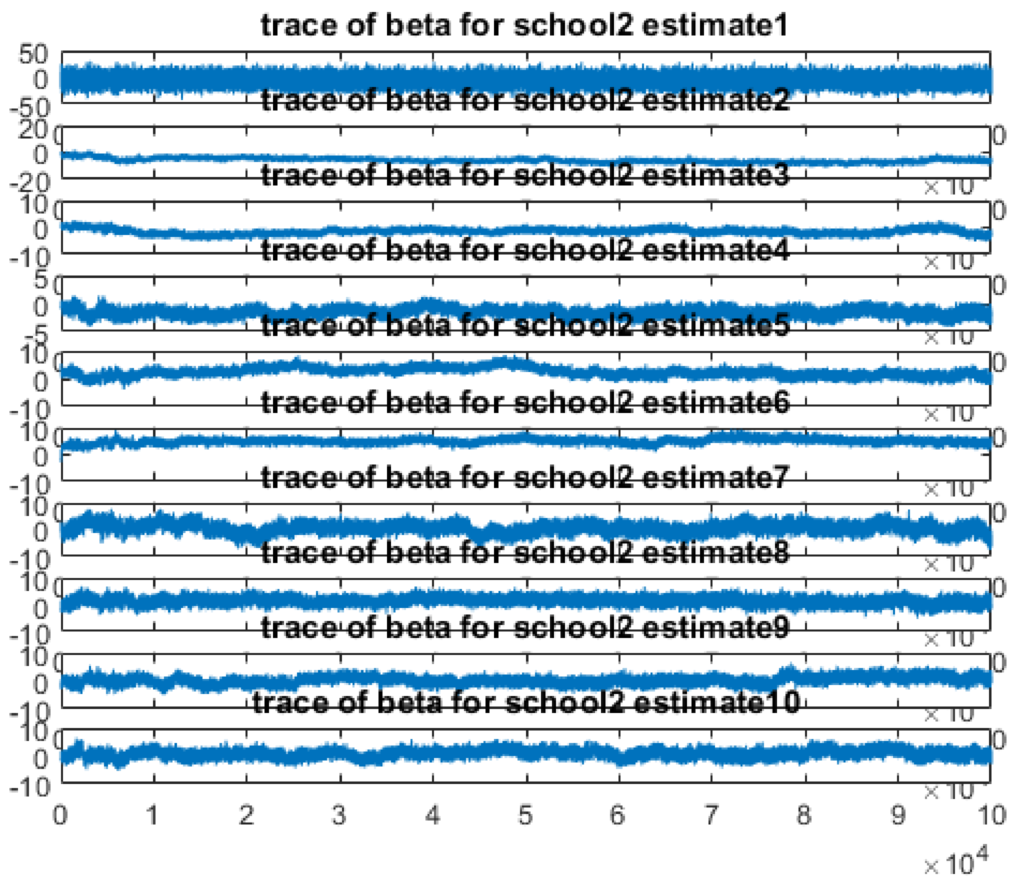

Figure B.10 Trace of coefficients for individual characteristics in high school 2, homogeneous sophistication

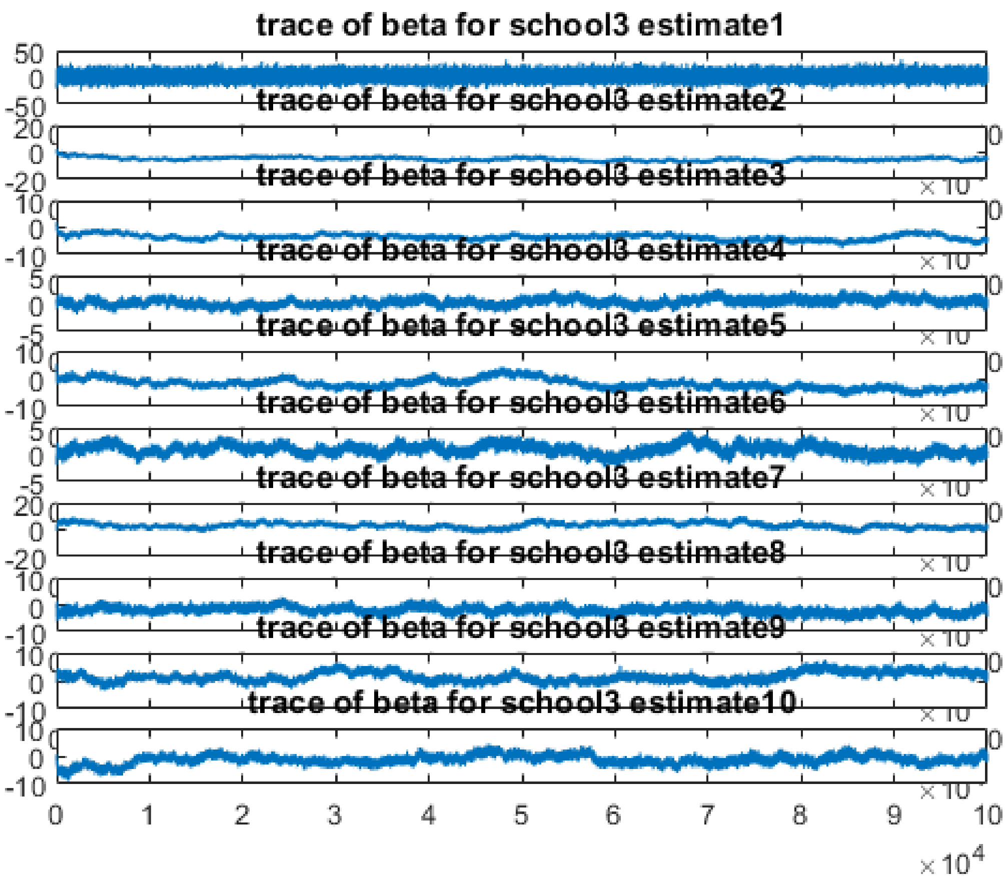

Figure B.11 Trace of coefficients for individual characteristics in high school 3, homogeneous sophistication

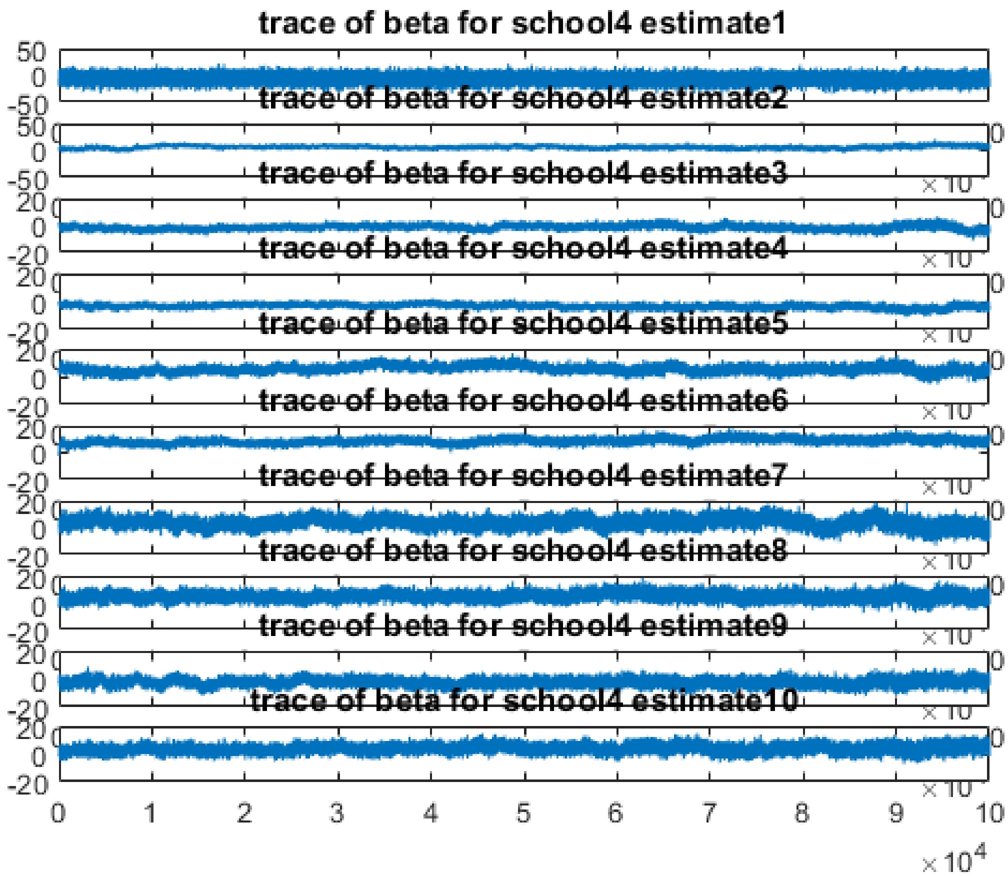

Figure B.12 Trace of coefficients for individual characteristics in high school 4, homogeneous sophistication

B.2 Heterogeneous Sophistication

For the same reason as homogeneous sophistication model, I would only show the plots for high school estimation here.

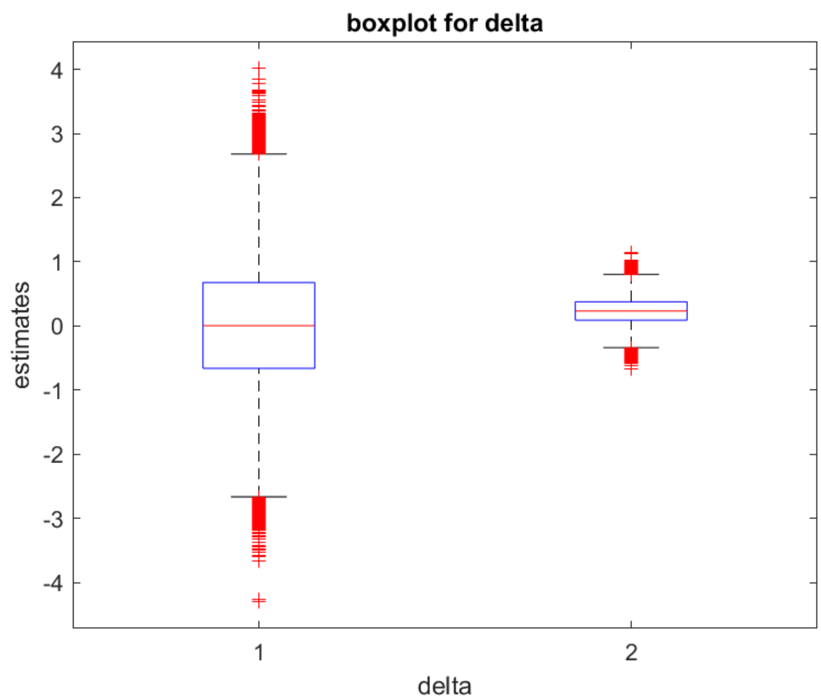

Figure B.13 Boxplot of coefficients for school characteristics, hetergeneous sophistication

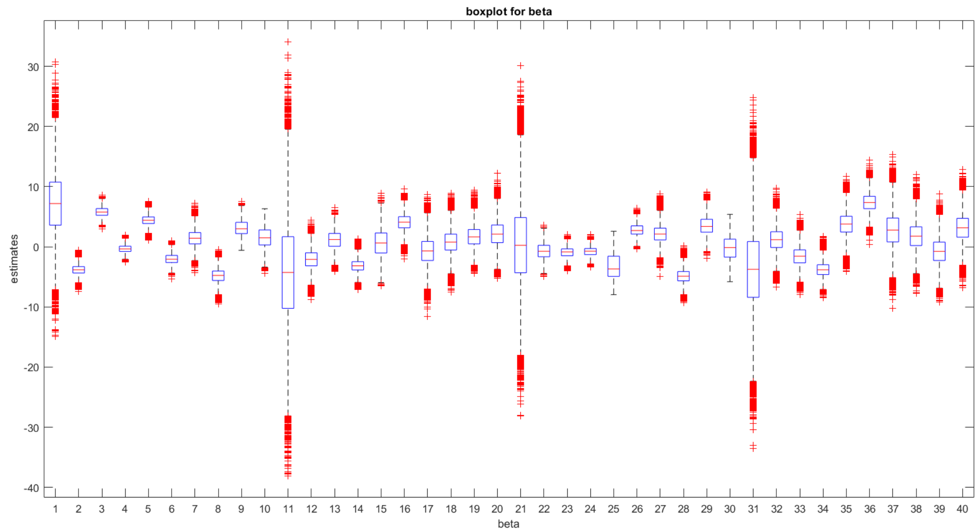

Figure B.14 Boxplot of coefficients for individual characteristics, hetergeneous sophistication

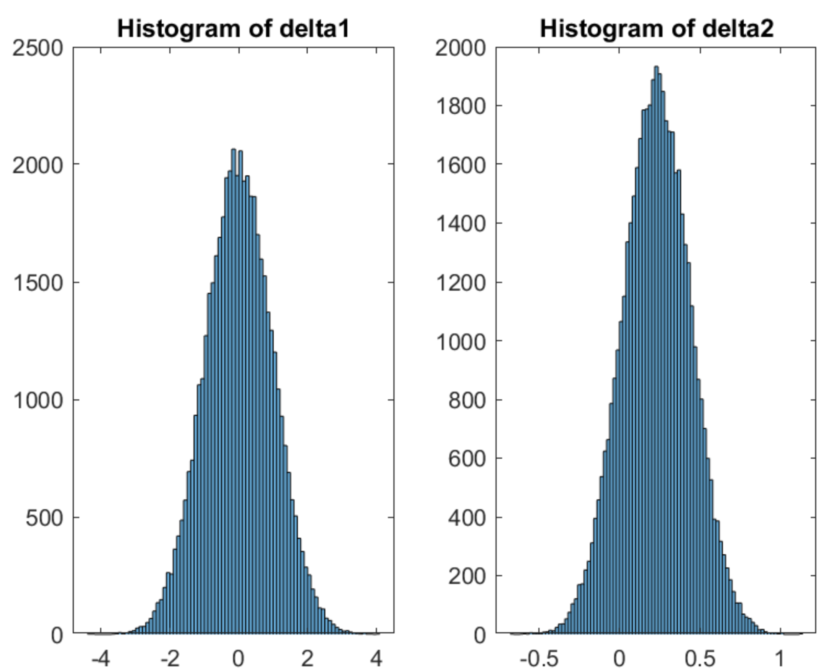

Figure B.15 Histogram of coefficients for school characteristics, hetergeneous sophistication

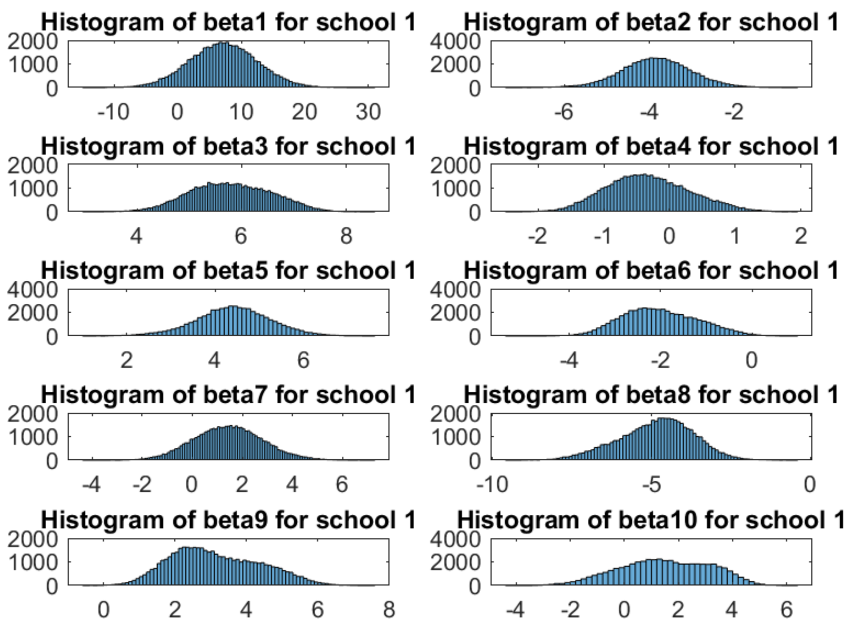

Figure B.16 Histogram of coefficients for individual characteristics in high school 1, hetergeneous sophistication

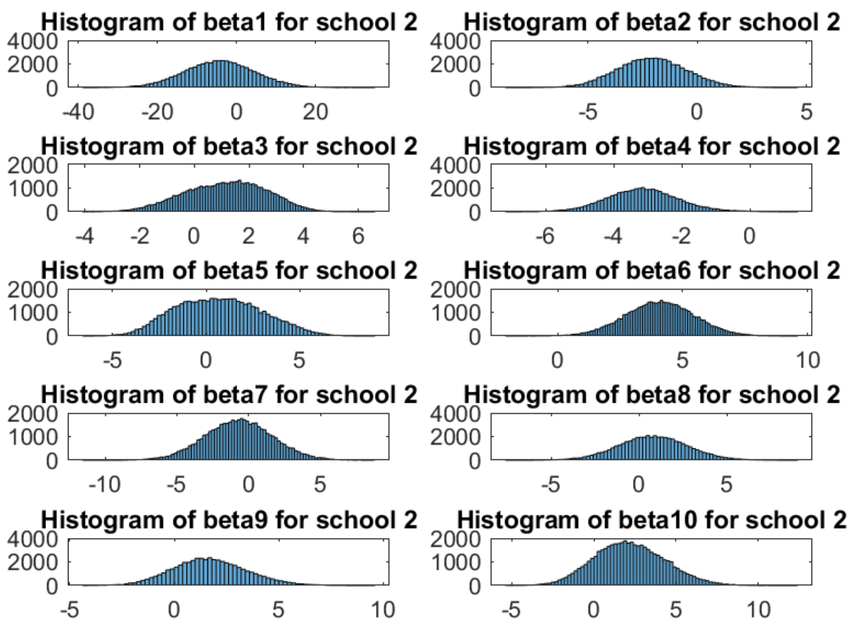

Figure B.17 Histogram of coefficients for individual characteristics in high school 2, hetergeneous sophistication

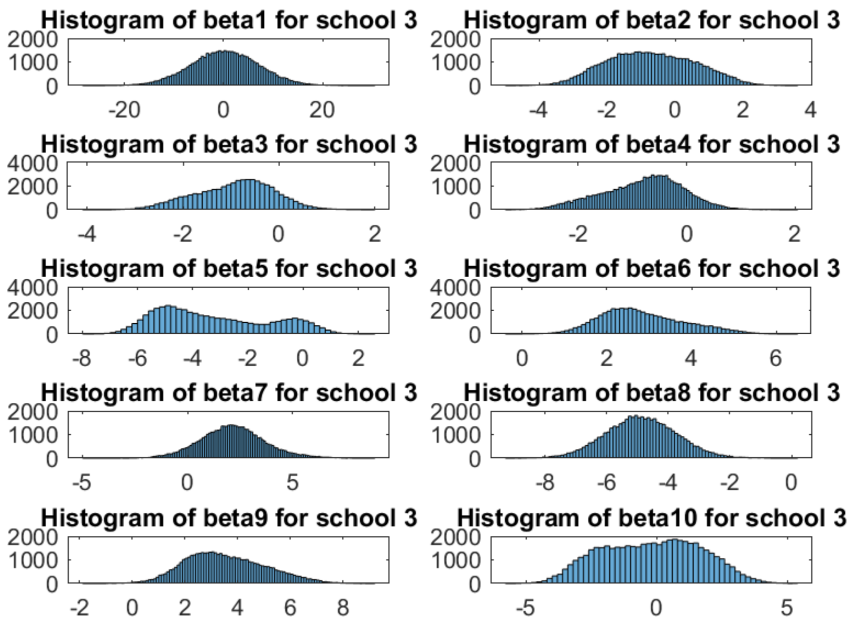

Figure B.18 Histogram of coefficients for individual characteristics in high school 3, hetergeneous sophistication

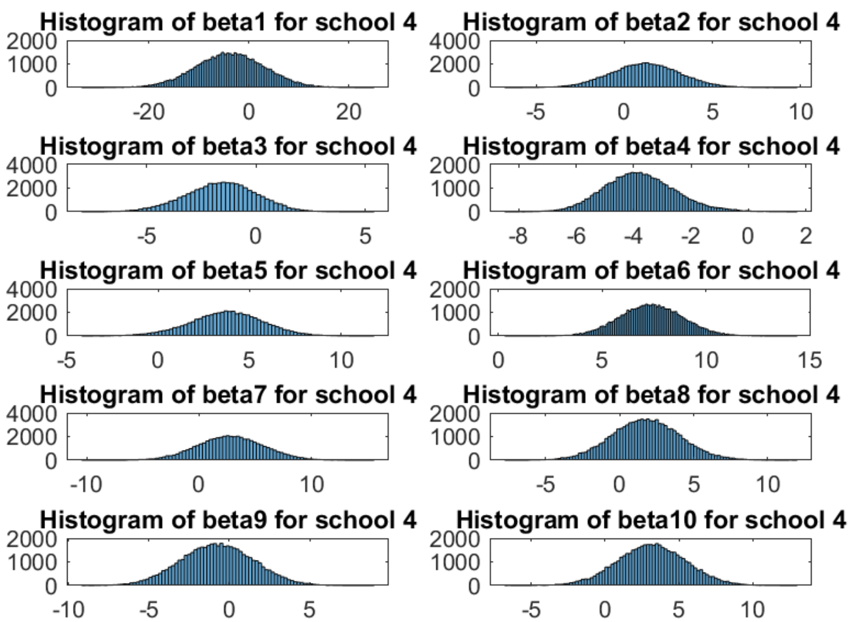

Figure B.19 Histogram of coefficients for individual characteristics in high school 4, hetergeneous sophistication

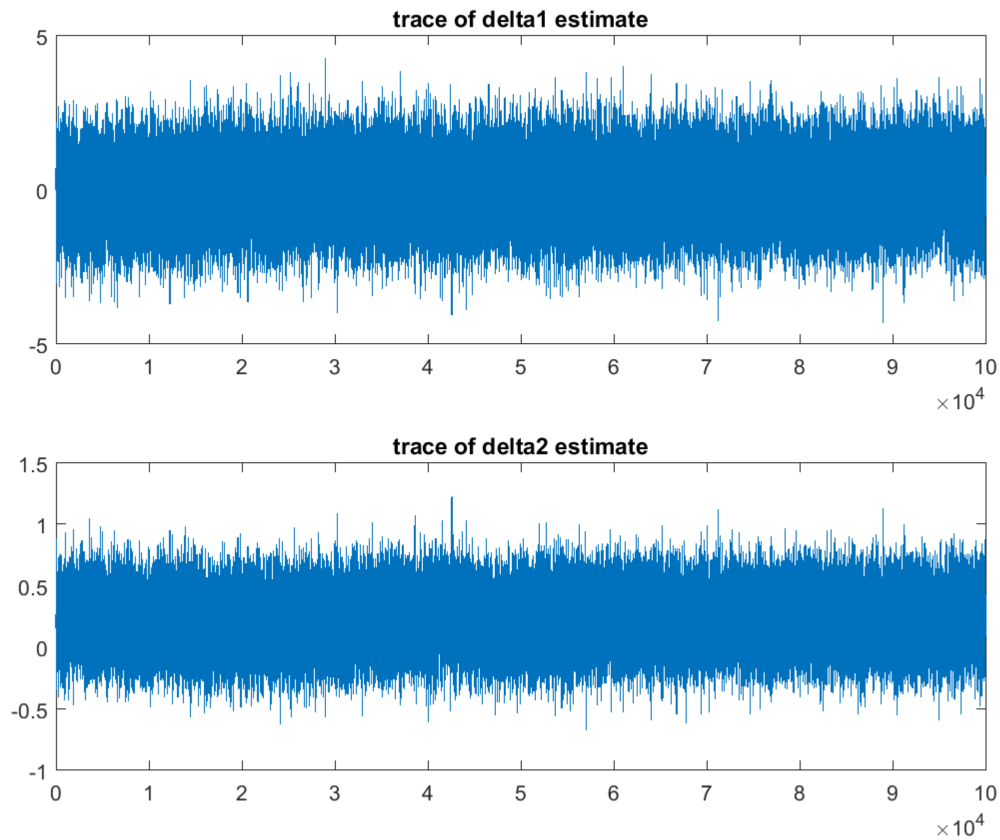

Figure B.20 Trace of coefficients for school characteristics, hetergeneous sophistication

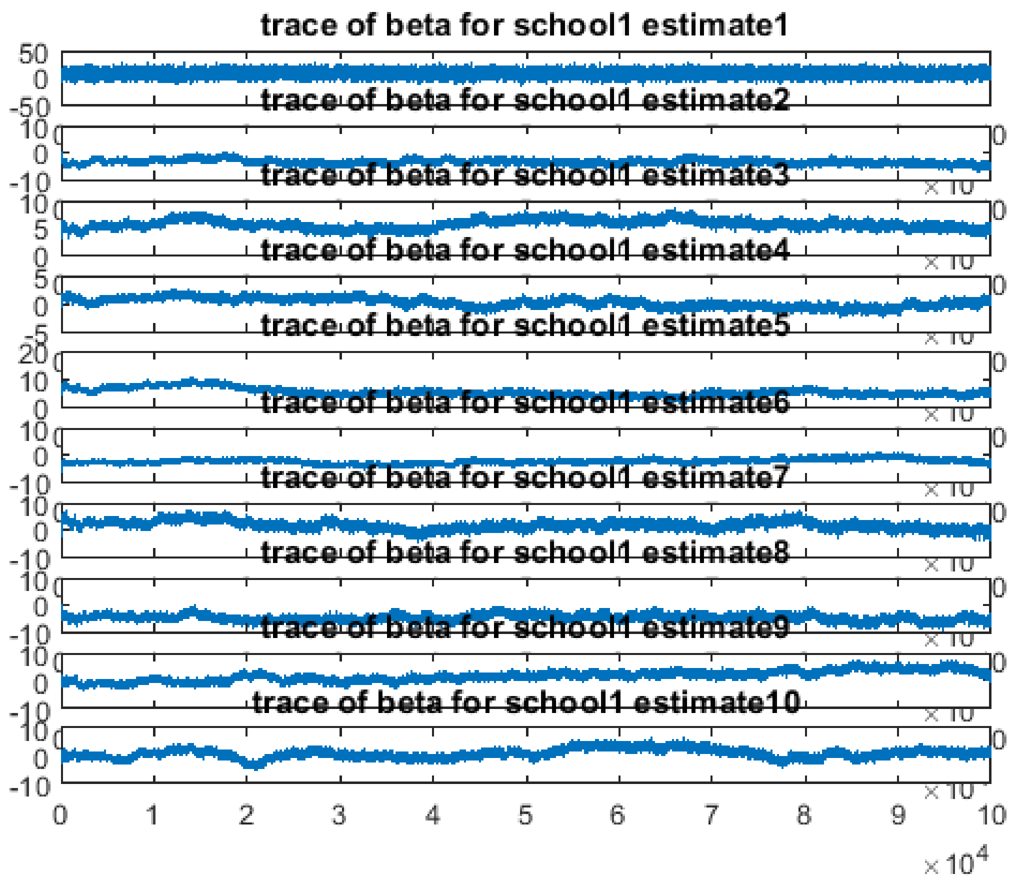

Figure B.21 Trace of coefficients for individual characteristics in high school 1, hetergeneous sophistication

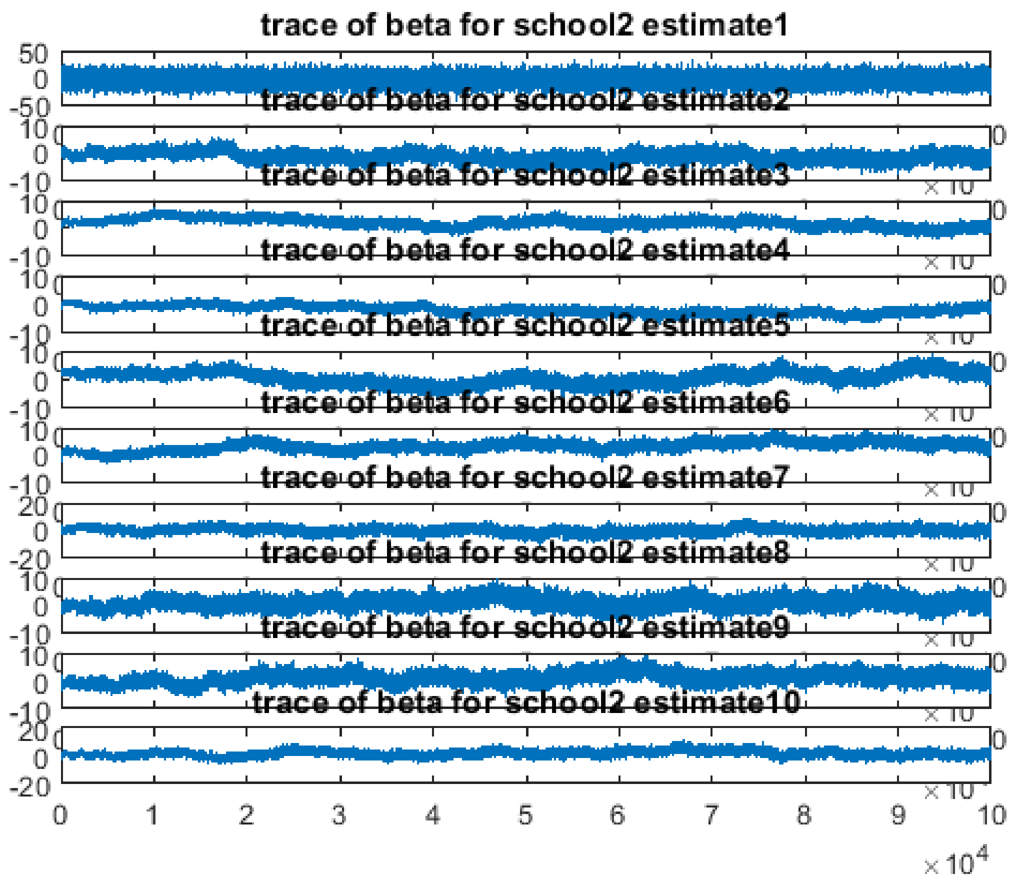

Figure B.22 Trace of coefficients for individual characteristics in high school 2, hetergeneous sophistication

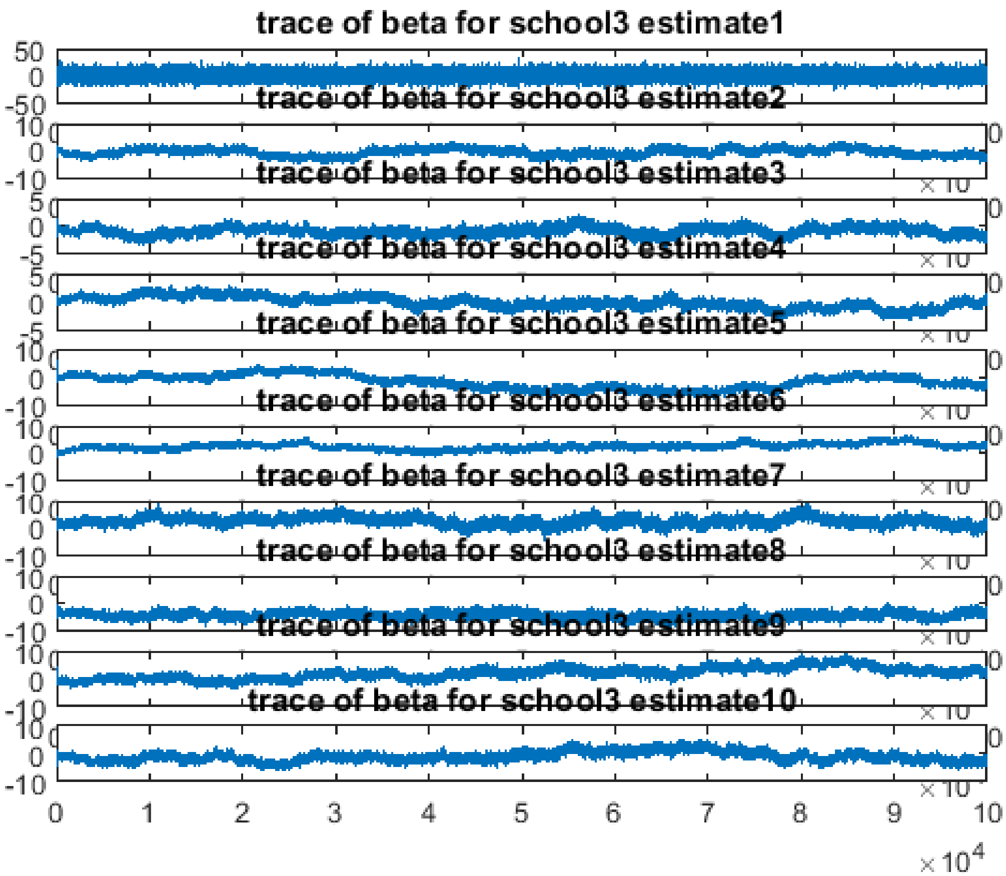

Figure B.23 Trace of coefficients for individual characteristics in high school 3, hetergeneous sophistication

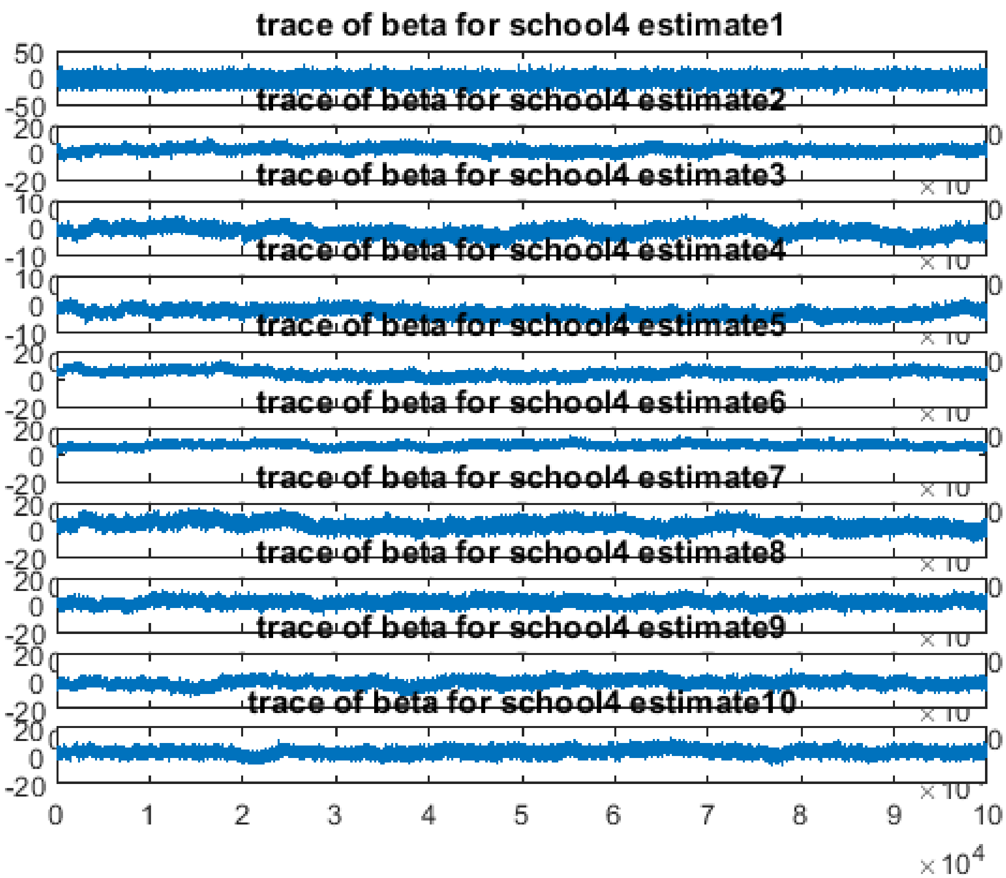

Figure B.24 Trace of coefficients for individual characteristics in high school 4, hetergeneous sophistication