An Empirical Analysis of School Choice under Uncertainty*

Kehinde Ajayi †   Modibo Sidibé ‡

April 12, 2017

Abstract

This paper develops a framework for quantifying student welfare in constrained school choice problems where admission chances are uncertain and interdependent across alternatives. We propose a strategy to construct an optimal portfolio of schools, that is appealing both theoretically and computationally, and when confronted with data allows us to recover individual preferences along with policy-invariant primitives that generate student choices. Our counterfactual simulations then investigate how uncertainty and constraints on the number of ranked choices impact student welfare. An application using administrative data from Ghana shows that expanding the number of ranked choices increases welfare and generates substantial redistribution across ability groups. Moreover, providing additional information that effectively reduces uncertainty delivers a more efficient allocation of students to schools, with comparable redistribution effects.

Keywords: Optimal portfolio, school choice, uncertainty

JEL Classification: C53, D61, I20

*We thank Atila Abdulkadiroglu, Attila Ambrus, Pat Bayer, Allan Collard-Wexler, Yingua He, Kevin Lang, Dilip Mookherjee, and participants at many seminars and conferences for their helpful comments and discussions. We also thank the Computerized School Selection and Placement System Secretariat, Ghana Education Service, and SISCO Ghana for providing data and background information. Ajayi is grateful for the hospitality of the Department of Economics at Duke University where a large part of this paper was completed.

†Boston University; kajayi@bu.edu

‡Duke University and CREST; modibo.sidibe@duke.edu

1 Introduction

Public school systems around the world are undergoing substantial reforms: several cities have adopted centralized coordinated assignment mechanisms to improve the matching between students and schools. Despite the prevalence of these systems, very little is known about the welfare consequences of the various practical aspects of school choice mechanism design.

Recent empirical research on school choice under centralized assignment focuses on the welfare gains associated with different assignment procedures.[1] Introducing student choices into a coherent demand framework for schools is a required first step toward welfare analysis but turns out to be a difficult undertaking. The underlying reason is that practical implementation concerns (such as incomplete information about the preferences of other agents and constraints in the length of rank ordered lists) prompt agents to strategize over their submitted list. As a consequence, estimates obtained under the assumption of truthful reporting of individual preferences may have limited validity.[2]

This paper bridges the gap between theoretical and empirical analyses of school choice. We provide a framework to estimate individual preferences and simulate the welfare effects of various policy alternatives in centralized assignment systems with uncertainty and a constrained number of choices. In our model, students endowed with a pure characteristics indirect utility, uncertain about their own type (individual test score), and the type of other agents (aggregate uncertainty) form expectations over subjective admission probabilities and solve a simultaneous search problem to construct an optimal portfolio of schools. The portfolio choice entails a complex large-scale combinatory optimization, which is NP-complete in general when the number of alternatives is large, and consequently no exact solution is available.

The essential features of our model, uncertainty and constrained choices, are present in most assignment systems, giving broad applicability to our framework.

[1] Ergin and Sonmez (2006) and Abdulkadiroglu et al. (2011) consider theoretical properties to compare alternative mechanisms, which are obtained under stringent conditions namely perfect information, unidimensional preferences, and unlimited number of ranked choices.

[2] For example, a popular assignment algorithm is Gale-Shapley's deferred acceptance (Gale and Shapley, 1962). It has the desirable property of incentivizing truth telling, but only when students can rank an unlimited number of choices. In practice, this algorithm is rarely adopted in systems that allow unlimited choices.

Uncertainty is a fundamental aspect of school choice – admission is rarely guaranteed but instead based on aggregate demand and individual priorities. Similarly, constraints on choices are a common feature of centralized coordinated assignment mechanisms.[3] It is therefore unsurprising that most applicants include some safety schools in their application set and do not merely apply to their favorite schools.[4]

The main methodological contribution of our paper is to propose a tractable strategy to construct an optimal portfolio of schools when individuals are uncertain about admission chances, which are interdependent across alternatives. A crucial assumption in our setting is that agents are sophisticated. The model is similar to the stochastic portfolio choice problem of Chade and Smith (2006), with the known difficulty of accounting for all substitution patterns induced by correlated admission chances. Learning from the failures of existing algorithms, we develop a strategy to economically expand the choice set, reintroducing the missing substitutions, and recover an optimal solution in a structure that is inherently sequential and tractable. Specifically, from an initial step where the optimal portfolio is easy to construct, we select a set of portfolios, that are close in value to the optimal one, to serve as candidates in the next step. Then, we iterate on those candidates until reaching the required portfolio size.

The generality of our approach to recover individual portfolio choices is not costless – we do not explicitly model individual formation of subjective admission probabilities. Instead, we follow the literature and assume that school admission is governed by cutoffs that can be imperfectly recovered using historical data. We allow for two-sided measurement error in these cutoffs that originates from over time variation in the preferences of other students (aggregate uncertainty), and limited information about a student's own test score (individual uncertainty). A student who receives a low score will have a low admission chance at all schools.

[3]Within the United States, school choice systems in Denver, Chicago, and New York limit the number of choices to 5, 6, and 12 respectively. Further afield, other systems also frequently impose limits such as for preschool choice in Barcelona (10), for secondary school applications in Ghana (4), Kenya (3), Trinidad and Tobago (4), and for college in Chile (8), Ontario (5), and Spain (8). Even in cases without explicit limits (e.g., college choice in Hungary), students often still face costs such as an application fee per listed choice.

[4]See Pallais (2015) for evidence of binding constraints in college application in the US and Chade et al. (2014) for an equilibrium analysis of schools and students decisions in a similar context.

The same interdependence in admission chances applies to lottery-based systems in which each student receives one lottery number that determines priority for admission to all schools.

We estimate this model using administrative data from Ghana's senior high school choice system in the year 2008, when students were allowed to rank a maximum of 6 programs. The assignment mechanism is a student-proposing deferred acceptance algorithm, with student priorities determined by scores on a standardized test. Our data consist of 128,468 students choosing between 1,933 programs.[5] Dimensionality is the major complication of the analysis, and since the outcome of the matching cannot be mapped when considering the full sample or a smaller subset of individuals, options are limited. Our empirical strategy, based on the method of simulated moments, estimates preference parameters that match the empirical characteristics of students' ranked choices to the ones predicted under the optimal portfolio choice model. Moments used in the estimation include summary characteristics of students as well as chosen programs.

We find that parameter estimates depend crucially on behavioral assumptions about how individuals construct a portfolio. Specifically, we compare our benchmark model of optimal portfolio choice given interdependent admission chances to two alternatives: i) truth telling, and ii) strategic portfolio choice assuming independent admission chances (as in Chade and Smith, 2006). Under the optimal portfolio choice strategy, parameter values are as expected indicating individuals' preferences for school quality and boarding facilities and a disutility associated with technical programs. Our benchmark model of strategic behavior generates a superior fit to a model that assumes truth telling, indicating that students are indeed likely to be strategic in response to restrictions on the number of choices they can list, and to alter their behavior as a result of interdependent admission chances. This comparison suggests that our optimal portfolio choice algorithm may provide valuable insight into school choice behavior under uncertainty, and that ignoring these common features of school choice systems may generate misleading predictions about the likely impacts of policy reforms. Our analysis of the model's fit also reveals an interesting pattern – the simulated portfolios under the optimal solution are substantially more selective than observed choices, which is

[5] By contrast the largest school district in the US is New York, which serves only 90,000 eligible students (Abdulkadiroglu et al., 2009).

not the case for the suboptimal, but easier to implement method of Chade and Smith (2006), suggesting that individuals may not be able to fully internalize the effects of interdependent admission probabilities.

Having validated the fit of our model, we use the model to understand the policy implications of school choice design by conducting two counterfactual experiments. Our first experiment investigates the relation between the number of ranked schools and welfare. Theoretically, the relationship between total welfare and the number of ranked choices is ambiguous. While a large number of ranked choices may increase overall match quality and hence total welfare, it prevents individuals from signaling the intensity of their preference, which may be ex-ante inefficient especially when admission chances are uncertain (Abdulkadiroglu et al., 2011). As a consequence, increasing the number of ranked choices produces winners and losers. The efficiency trade-off may be even more complex when students from disadvantaged backgrounds are unable to signal their preferences because of a larger exposure to uncertainty. To quantify the importance of these mechanisms, we use our estimated preferences to simulate optimal portfolios for each individual while varying the number of ranked choices. Then, we run the student-proposing deferred acceptance matching algorithm to determine admission outcomes, and impute total welfare.

We find that there are substantial welfare improvements from allowing students to rank an unlimited number of choices. Total welfare is a concave function of the number of choices permitted – expanding the number of choices from 1 to 2 increases total welfare by 26.4% and reduces the share of unassigned students from 72 to 55 percent. Allowing students to submit an unrestricted number of choices generates approximately 2.4 times as much total welfare as allowing for a single choice. As expected, increasing the number of choices produces winners and losers – a redistribution from high performing to low performing students ensues. While the 10% lowest ability students make up only 3.1% of total welfare under a single choice, they constitute 8.7% when students are allowed to submit an unlimited number of choices. Similarly, the 10% highest ability students comprise 33.7% of total welfare with a single choice, compared to 11.8% with an unlimited number of choices.

Our second experiment is motivated by the observation that although our model fits the data well, observed portfolios are less selective than the optimal ones, which is not surprising given the complexity of the problem. We propose a realistic policy intervention that removes individual uncertainty and implies a simple strategy to construct application portfolios. Specifically, we provide students with their test score, and a test score cutoff for admission to each school (based on admissions in the previous year). Students can apply to a school only if their test score exceeds the cutoff. We find that there are substantial potential welfare improvements from providing this additional information – total welfare increases by 38% when only one choice is allowed, and providing six choices almost eliminates administrative assignment. Once again, lower ability students experience the largest welfare gains.

We extend the existing literature in several directions. First, we introduce a novel approach to model the decision problem of students, incompletely informed about the characteristics of other participants in the matching, when faced with a large set of interdependent choices. Our strategy allows us to handle cases with large numbers of both students and schools by explicitly modeling the initial portfolio choice process.

Second, we directly assess the welfare implications of constraints on choices. Most school choice reforms are motivated by the fact that constraining the number of choices and forcing students to strategize is costly. The traditional approach to comparing the efficiency of various assignment mechanisms consists of running contextual lab experiments (Chen and Sonmez, 2006; Calsamiglia et al., 2010; Troyan, 2012). By contrast, our policy simulations allow for an analysis of marketwide changes using data on the observed behavior of agents in a large school choice system.

Despite the large literature studying individual preferences in centralized assignment systems, until very recently little attention has been given to the strategic content of reported rankings of various school options. Existing empirical studies either identify preferences using rank orderings within a submitted portfolio, taking the selected portfolio as given (Ajayi, 2013; Abdulkadiroglu et al., 2015) or consider a narrow set of applicable mechanisms (He, 2012; Calsamiglia et al., 2014; Agarwal and Somaini, 2014) and a limited set of policy implications (Fack et al., 2015).

To illustrate the innovation in our approach, we elaborate on how our work compares to other recent papers addressing strategic behavior. The predominant focus in the empirical school choice literature has been on lottery-based admission and the Boston mechanism.[6] He (2012) considers school choice in a neighborhood of Beijing, where students apply for admission to one of 4 local middle schools and are assigned using random lotteries and the Boston mechanism. Given the small number of available choices, his estimation approach relies on being able to explicitly specify the full set of portfolio alternatives, which is computationally infeasible in larger markets. Calsamiglia et al. (2014) exploit the main strategic implications of the Boston mechanism to derive feasible choice sets for each round of the portfolio choice. Agarwal and Somaini (2014) propose an estimation strategy that applies to a larger class of mechanisms but which nonetheless maintains the condition that assignment is based on coarse priority types with lotteries used to break ties. Hence, their methodology does not apply to cases with merit-based admission. In contrast to these three papers, we provide a more general approach to estimating student preferences.

The closest paper to ours in terms of its applicability is Fack et al. (2015). Broadening their scope of analysis, the authors study a case of merit-based admission under the student-proposing deferred acceptance algorithm, that is similar to our application. While the paper proposes a menu of approaches to estimate agents' preferences under a range of assumptions, the authors do not explicitly model the portfolio choice process and are thus limited in their ability to perform counterfactual policy experiments, such as relaxing constraints on choice lists.

In sum, the central purpose of our paper is to provide a tractable model of optimal portfolio choice for large matching markets that allows us to generate new insights about the welfare implications of two fundamental features of school choice mechanism design (uncertainty and constraints on choices); we estimate both agents' preferences and their subjective beliefs about their admission chances; and we use our parameters to simulate counterfactual policies.

The paper proceeds as follows. Section 2 describes our context and the data generated by the senior high school admission system in Ghana. We illustrate the key features of our problem with a stylized example in Section 3. Section 4 introduces our model of optimal portfolio choice and describes our solution concept. Section 5 outlines our parametric assumptions and estimation strategy.

[6] Also known as an immediate acceptance algorithm, this mechanism was used for admission to Boston Public Schools between July 1999 and July 2005 (see Abdulkadiroglu and Sonmez, 2003).

Sections 6 and 7 present our results and policy simulations. Finally, we conclude.

2 Senior High School Choice in Ghana

2.1 Background

The national school system in Ghana consists of six years of primary school, three years of junior high school (JHS), and three years of senior high school (SHS). Duflo et al. (2017) report that under 20% of SHS graduates enroll in tertiary education directly after senior high school.[7] Students completing junior high school apply for admission to senior high school through a centralized application system. Students apply to specific academic programs within a school and can submit a ranked list of up to four programs. Available programs include agriculture, business, general arts, general science, home economics, technical studies, visual arts, and several occupational programs offered by technical or vocational institutes. After submitting their ranked lists of choices, students take a standardized Basic Education Certification Exam (BECE). The application system then allocates students to schools based on their BECE scores and a deferred acceptance assignment algorithm described in more detail when we discuss our policy experiments in Section 7.

As in many other coordinated school choice systems, students in Ghana are uncertain about their admission chances when they select their choices. Students apply before taking the BECE so they do not know their exam scores when they submit their ranked lists. Moreover, schools do not specify the required scores for admission but instead only report the number of vacancies available in each academic program they offer. Admission cutoffs are therefore endogenously determined by the distribution of vacancies and application choices in a given year.

2.2 Data

We use administrative data on the universe of senior high school applicants in Ghana's centralized school choice system for our empirical analysis. The data

[7] This percentage may be lower than average as their sample consists of a relatively low-income population – people who didn't enroll in SHS immediately after JHS because they couldn't initially afford it.

include students' exam scores, their ranked list of chosen programs, their admission outcomes, and basic demographics (gender, age, and the junior high school attended). Our primary estimation sample focuses on the 2008 cohort of senior high school admits. Students could rank up to six choices from the 1,933 programs available in over 500 schools that year. We have information on 128,468 students who submitted a complete list of six programs ranked in order of preference.

Table 1: Student Characteristics

	Mean	Median	Std. Dev.	Min.	Max.	Obs.
Male	0.579	1.000	0.494	0	1	128468
Age	16.609	16.000	1.686	9	54	128468
BECE exam score	292.900	285.000	50.485	185	469	128468
Mean BECE in JHS	292.304	281.812	40.202	199	432	128468
Attended public JHS	0.757	1.000	0.429	0	1	128468
Admitted to first choice	0.268	0.000	0.443	0	1	128468
Admitted to second choice	0.204	0.000	0.403	0	1	128468
Admitted to third choice	0.188	0.000	0.390	0	1	128468
Admitted to fourth choice	0.166	0.000	0.372	0	1	128468
Admitted to fifth choice	0.026	0.000	0.158	0	1	128468
Admitted to sixth choice	0.019	0.000	0.137	0	1	128468
Administratively assigned	0.130	0.000	0.337	0	1	128468

Table 1 summarizes student demographics and admission outcomes. Over half of the students are male and the median age is 16. Student performance on the BECE exam ranges from 185 to 469 points out of a possible 600, so students have very different chances of gaining admission to any given program. We do not have information on family background, so we use junior high school characteristics as our proxy for students' socio-economic status. Junior high schools vary considerably in their average performance and 76 percent of students attended a public junior high school. Almost 27 percent of students were admitted to their first choice program, while less than 2 percent were admitted to their sixth choice. A sizable 13 percent of students were rejected from all six of their chosen schools and administratively assigned to an undersubscribed program at the end of the assignment process.

Table 2 summarizes school and program characteristics. There is substantial variation across programs. As in Abdulkadiroglu et al. (2014) and Pop-Eleches

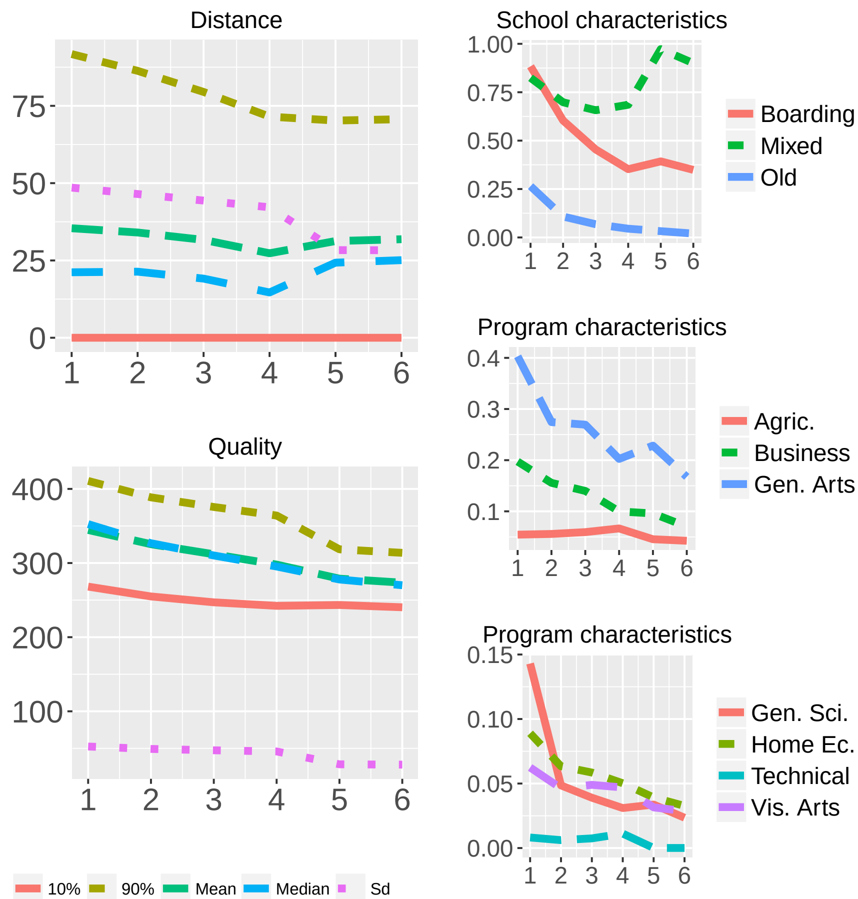
Table 2: Program Characteristics

	Mean	Median	Std. Dev.	Min.	Max.	Obs.
Mean BECE of admits	285.528	269.155	51.616	180	446	1933
Number of admits	63.071	49.000	47.187	0	358	1933
Programs offered	4.845	5.000	1.314	1	14	1933
Public	0.998	1.000	0.039	0	1	1933
Boarding facilities	0.602	1.000	0.490	0	1	1933
Pre-independence	0.088	0.000	0.284	0	1	1933
Technical/vocational inst.	0.013	0.000	0.113	0	1	1933
Agriculture	0.162	0.000	0.369	0	1	1933
Business	0.192	0.000	0.394	0	1	1933
General Arts	0.235	0.000	0.424	0	1	1933
General Science	0.130	0.000	0.336	0	1	1933
Home Economics	0.191	0.000	0.394	0	1	1933
Technical Studies	0.070	0.000	0.256	0	1	1933
Visual Arts	0.126	0.000	0.332	0	1	1933

and Urquiola (2013), we measure school quality as the average exam score of students admitted to each program in the previous year, which ranged from 180 to 446 points, with a mean of 286 and standard deviation of 52 points. The average program admitted 63 students but this ranged from 0 to 358. Each school offered an average of 5 programs, with some offering as many as 14 different programs. 99.8 percent of programs were offered by public schools, and 60 percent of them were in schools with boarding facilities. Only 9 percent of programs were offered by schools established by the British colonial administration before Ghana gained independence in 1957. General arts was the most commonly offered program, accounting for a quarter of available choices.

Figure 1 presents descriptive statistics on students' ranked program choices. We begin by examining the distance between a student's junior high school and selected senior high school. We do not have exact coordinates for school locations so we measure the distance between centroids of the 110 administrative districts in the country. Ghana's school choice system is truly national and some students apply to schools as far as 450 miles away (roughly the distance from Boston to Washington, DC). Preferences for distance are convex. Students' first choice programs are on average 35.1 miles away from their junior high schools

Figure 1: Distribution of variables



Note: Left two panels illustrate distribution of distance and quality of selected choices. Right three panels plot hazard rates for selecting a program with a given characteristic.

and their second choice programs are 1.3 miles closer to them. Their third and fourth ranked choices are 31.4 and 27.1 miles away, but their last two choices are further away at a distance of 31.1 and 31.7 miles on average. Even though there is no clear gradient, the dispersion in distance decreases over the six choices with standard deviations of 48.54, 46.48, 44.37, 42.19, 28.35, and 28.30, respectively.

In contrast to preferences for distance, peer quality in ranked programs decreases monotonically. The average exam score of a students' first choice program is 343 and this falls to 273 for the lowest ranked choice. This is a difference of 1.2 standard deviations in the peer quality distribution. Considering preferences for distance together with preferences for academic quality, it appears that students are willing to travel for the opportunity to attend a high quality program but less willing to travel for their lower ranked, lower quality choices.

The last three panels in Figure 1 examine discrete program characteristics and reveal more patterns in aggregate choices. Each line plots a hazard rate indicating the probability that a student lists a program with a particular characteristic given that they have not listed a program with that characteristic for a previously listed choice. Students prefer programs in boarding schools and have an increasing likelihood of selecting mixed schools but a decreasing likelihood of selecting older schools established before Ghana gained independence – 88 percent of students choose a program in a boarding school as their first choice and only 59 percent select one as their lowest ranked choice; 83 percent of students choose programs in mixed sex schools as a first choice but 98 percent do for their sixth choice; 27 percent of students choose programs in schools constructed before independence as a first choice but only 2 percent do for a sixth choice.

Finally, we illustrate students' preferences over academic program tracks. General arts is the most popular program track, with 40 percent of students choosing this program as their first choice and 43 percent choosing a general arts program as their sixth choice. General science has the steepest gradient in choices. 14 percent of students choose a general science program as their first choice and only 7 percent choose one as their sixth choice. Preferences for agriculture programs show the reverse pattern, with 6 percent of students choosing one as their first choice and 9 percent choosing one as their sixth choice. The remaining programs are relatively equally represented across choices with an average of 20 percent of students choosing business programs, 10 percent choosing home economics programs, 7 percent choosing visual arts programs, and 4 percent choosing technical programs.

3 Problem

This section illustrates the problem considered in this paper. For expositional clarity, we propose an extreme simplification, but still a useful description. We consider the matching problem of three students $1, 2$, and 3 to three schools A, B, and C with one vacancy each. Without loss of generality, we assume that there is no idiosyncrasy and schools give the same utility to every student. We maintain the assumption that allocation is based on a student-proposing deferred acceptance algorithm.[8]

Table 3: Matching table

Students	1	2	3	Schools	A	B	C
Test score	10	12	9	Utility	7.9	10	11.4

Under perfect or imperfect information, and for a given priority structure, individual choice is easily constructed. In that setting, the length of the rank ordered list is not important, as strategic considerations by agents yield an efficient equilibrium allocation – explicitly, students apply to the single school with highest utility given their beliefs about other students' behavior. To see the roles of constraints in the number of choices, it is important to consider a setting where individuals have incomplete information. Specifically, when students are uncertain about their priority (test score or ranking in the distribution of test scores), the optimal strategy consists of reporting true preferences when there is no constraint on the length of the reported list. However, most allocation mechanisms include a constraint on the length of the rank ordered list. In our current case, allowing students to list only two choices will leave student 3 unmatched (and thus administratively assigned) if agents report their true preferences and priorities are based test score. As a consequence, constructing an optimal list of schools that

[8] Relaxing this assumption in this illustration is trivial, but more complex in a more realistic example. We discuss this issue further in section 7.

balances student ambition to get into the best available school with an insurance against administrative assignment is the goal of the next section.

4 Model

We develop a framework to understand the application behavior of students to schools. Each student submits an ordered list of choices, after which a planner assigns students using the deferred acceptance mechanism according to their application list, the available capacity of each program, and a predefined priority.

4.1 Definition

A finite set of students $\mathcal{I} = \{1, 2, \ldots, I\}$ apply to a finite set of schools $\mathcal{J} = \{1, 2, \ldots, J\}$.[9] Each school has positive capacity, and students can opt out of the matching system and enjoy an outside utility u_0, which we set to 0 for simplicity. A student is characterized by a set of observed attributes X_i and a test score ψ_i which is unknown when they submit their choices. The latter defines individual admission priorities while the former captures his preferences. Schools have an observable set of characteristics given by Z_j, and a fixed capacity denoted by C_j. Since all students in our sample submit the required number of choices, we assume an infinite cost for exceeding the constraint but that there is no application cost otherwise. The utility for an individual i with characteristics X_i matched with a school with attributes Z_j is given by $U(X_i, Z_j)$. We follow Berry and Pakes (2007), and assume that the indirect utility function includes a disturbance term ϵ that is additively separable from observable characteristics (school attributes Z and student characteristics X).

$$U_{ij} = \gamma Z_j - d(l_i, l_j) + \sum_{k=1}^{K} \Gamma_k Z_j^k X_i^k + \epsilon_{ij} \tag{1}$$

where the set of school attributes, Z_j, includes quality, size, and indicators for boarding facilities, old, and program track. These characteristics are summarized in Table 2. The set of individual characteristics, X_i, consists of individual test

[9] In our empirical application, we define a choice as a bundle (school,program). We have 517 schools, each offering between 1 and 14 academic programs, which yields a total of 1,933 choices.

score, gender, age, and middle school average test score, and $K = \mathbf{dim}(X) \times \mathbf{dim}(Z)$.[10] These characteristics are described in Table 1. And $d()$ provides the distance between student i's location l_i and school j's location, l_j. In addition, we include interaction terms between distance, school characteristics, and individual attributes. Finally, ϵ_{ij} is an idiosyncratic error term, with $\varepsilon_{ij} \sim \mathcal{N}(0, \varsigma)$. Since over 99 percent of programs are public schools, we use distance as our numeraire, and measure utilities in terms of willingness to travel.

Under the assumption that students are price takers, we can treat the construction of a portfolio of schools at the individual level. Denoting by q_{in} the admission chance for an individual with test score ψ_i at school with capacity C_n, the optimal portfolio of an individual i, of size $N = \|\mathcal{S}\|$, is denoted by $\mathcal{S}_i$ and satisfies:

$$\sup_{\mathcal{S}_i} f(\mathcal{S}_i) = \sum_{n=1}^{N} \mathcal{P}(q_{in}|q_{i1}, \ldots, q_{i,n-1}; X_{-i}) U_{in} \tag{2}$$

where $\mathcal{P}$ is the probability that student i is admitted to choice n, given his rejection from all other higher ranked choices $(1, \ldots, n-1)$, and the preferences of other agents X_i. Obviously when $N = J$, admission chances do not play any role, and the problem simplifies to submitting an ordering of schools by utility,

Finally, to close the model, we describe how student i forms beliefs about his admission chance at school j. There are two sources of uncertainty in our model. The first, which we refer to as individual uncertainty, comes from the fact that individuals apply to schools before taking the exam that determines their ranking in the matching algorithm. The second, which we refer to as aggregate uncertainty, comes from limited information on the characteristics of other market participants. We introduce a cutoff structure that allows us to capture these elements.

Definition Let Ψ_j be a cutoff : the minimal test-score required for admission at school j. There exists a unique vector of market clearing cutoffs $\widehat{\Psi}$ such that

$$\mathcal{D}_j(\widehat{\Psi_j}) \leq C_j \quad \forall j \in \mathcal{J} \tag{3}$$

Where $\mathcal{D}_j$ is the aggregate demand for school j. The former definition is the

[10] In a previous version, we include additional attributes. We focus in this version on the main sources of variation in the data.

standard framework in large matching models (see e.g., Azevedo and Leshno, 2016). We further assume that students can imperfectly forecast these cutoffs using historical data.[11] Yearly heterogeneity in the composition of the student body and variation in the school level capacities create uncertainty such that optimal cutoffs $\widehat{\Psi}_j$ are given by:

$$\widehat{\Psi}_j = \Psi_j + \zeta_j \tag{4}$$

where ζ_j is an error term, which is left unrestricted for now, and Ψ_j is the cutoff observed in previous year. Similarly, since students submit school applications prior to taking the exam, the real test score ξ_i is measured with error and given by:

$$\widehat{\psi}_i = \psi_i + \xi_i \tag{5}$$

Given the former definitions, the admission chance for individual i applying to school j can be written as:

$$q_{ij} = P(\widehat{\psi}_i > wj) = P(\psi_i - \Psi_j > \eta_{ij}) = 1 - F(\psi_i - \Psi_j) \tag{6}$$

where $\eta_{ij} = \zeta_j - \xi_i$ with cumulative distribution function $F(.)$. Since a choice is observed only when an individual i selects a school j, we can not separately identify the components of the errors associated with the measurement of individual test scores and cutoffs. As a consequence, we do not take a stand on the respective contribution of individual and aggregate uncertainty. Because of the error term, and in contrast to the standard literature, individuals observe a probabilistic signal of admission chance which is not a dummy variable.

In the empirical application, we assume that $\eta_{ij} \sim \Phi(0, \sigma_i)$ with the standard deviation of the error terms parametrized as follows:

$$\sigma_i = \exp(\sigma_0 + \sigma_1 \psi_i + \sigma_2 \psi_m) \tag{7}$$

where ψ_i is the individual test score, and ψ_m the quality of the middle school as measured by the average test score of students. Under this parametrization, we hope to account for the fact that high achieving students may face less uncertainty

[11] See Pathak and Shi (2014) for an empirical validation of this strategy.

regarding their performance on the final exam. Similarly, students from high quality middle schools are more likely to have teachers or parents who help them decide which schools to apply to. This strategy allows us to be agnostic about the content of agents' information sets at the time of the school choice decision as well as about the explicit model that generates cutoffs.

Importantly, given this admission probability, the conditional probability of getting admitted to a school with admission probability q_{in} when the $(n-1)$ bids have been unsuccessful is given by:

$$\begin{aligned}\mathcal{P}(q_{in}|q_{i1},\ldots,q_{i,n-1}) &= \mathcal{P}(\psi_i - \Psi_n > \eta_{in}|\psi_i - \Psi_1 < \eta_{i1},\ldots,\psi_i - \Psi_{n-1} < \eta_{in-1}) \\ &= F(\psi_i - \Psi_{n-1}) - F(\psi_i - \Psi_n) \end{aligned} \tag{8}$$

Note that $\mathcal{P}(q_{in}|q_{i1},\ldots,q_{i,n-1}) = \mathcal{P}(q_{in}|q_{i,n-1})$.[12]

4.2 Solution Method

This section considers the computation of an optimal solution to the problem described above. In the first part, we review existing methods and document their limitations. Then, we describe our solution method.

4.2.1 Existing Approach

The problem of constructing an optimal portfolio of schools has been analyzed in the literature, notably by Chade and Smith (2006), who proposed the Marginal Improvement Algorithm (MIA). The idea is to construct the global optimum by sequentially iterating on a local optimum. Starting from a first local optimum based on maximizing expected utility, additional optima are sequentially selected using the choice that yields the highest marginal improvement from an initial portfolio. The algorithm is detailed in Appendix A.1. Unfortunately, the algorithm does not reach the optimal solution when there is interdependency between admission chances. When admission chances are correlated (e.g., one test score used for admission at all schools – as is the case in Ghana and many other merit-based admission settings), additional substitution between schools implies that schools not

[12]Any extension of our model to deal with other allocation systems such as the Boston Mechanism would require a strategy to derive the admission chance probabilities, and the latter property of conditional probabilities will most certainly be lost. We leave this extension for future research.

selected in an earlier round might be part of an optimal portfolio in a later round. This failure can be illustrated using our running example. Suppose student 1 faced the following admission probabilities.

Table 4: Interdependent school choice

Schools	A	B	C
Utility	7.9	10	11.4
Adm Prob	1	0.8	0.7

Given the utilities and admission probabilities, MIA would begin by selecting school B as the single choice with the highest expected utility of 8 (relative to 7.9 and 7.98 for A and C). The next step would compare the marginal improvement of adding A or C – effectively comparing $\{B, A\} = 8 + (1 - 0.8)(7.9) = 9.58$ to $\{C, B\} = 7.98 + (0.8 - 0.7)(10) = 8.98$. This comparison would incorrectly identify $\{B, A\}$ as the optimal portfolio when in fact the optimal portfolio in the case of two permitted choices would be $\{C, A\}$, generating an expected utility of $7.98 + (1 - 0.7)(7.9) = 10.35$. The key reason why MIA fails is that rejection from a chosen school affects expected admission chances in subsequently listed choices.[13]

Instead consider an initial step that evaluates all possible options, then apply MIA to all of these initial portfolios and keep iterating. We would in this case initiate MIA by considering the set of three portfolios $\{A\}$, $\{B\}$, and $\{C\}$, and proceed by finding the marginal improvement to each, yielding three subsequent options – $\{C, A\}$, $\{B, A\}$, and $\{C, A\}$. In the final step, selecting between these three options would correctly yield $\{C, A\}$ as the optimal one. Because we only have three schools in this illustrative example, there is no difference between this alternative approach and the exact solution. However, when the number of alternatives increases, this approach would identify the optimal portfolio with fewer calculations than required to evaluate every possible combination. We describe our approach more formally in the next section.

[13] To be explicit, the conditional probability of admission to choice 2 given rejection from choice 1 is $q_{i2} - q_{i1}$ for interdependent admission chances and $(1 - q_{i1})q_{i2}$ for independent admission chances. Thus, the expected utilities given independent admission chances would be 9.58, 10.38, and 10.35 for $\{B, A\}$, $\{C, B\}$, and $\{C, A\}$.

4.2.2 Our Approach

Our problem belongs to the general family of Bandit Problems with Dependent Arms, which is *NP-complete* when the number of arms is large (Laporte, 1992). The *NP-completeness* comes with failure of all prominent algorithms, and the impossibility of designing an efficient algorithm. The underlying explanation is related to the same difficulties that require assumptions such as the gross-substitutes assumption in the matching literature (Kelso and Crawford, 1982).[14] In the presence of dependency between alternatives, properties such as the downward-recursive structure of the optimal solution are lost. As a consequence, any greedy algorithm will fail to produce an optimal solution. In addition, standard combinatory analysis, which requires computing all alternative portfolios, and selecting the one with the highest utility is impraticable when the number of alternatives is large: choosing 6 schools out of $1,933$ requires evaluating a total of $\binom{1,933}{6}$ alternative portfolios for each individual.

To overcome these limitations, we propose an approximation method. The solution is based on the premise that the problem is inherently sequential, and we exploit its structure while overcoming the limitations of simple greedy algorithms.

Sub Modular Improvement Algorithm (SMIA): Solve $\binom{N}{J}$

Step 1: Set $L < J$, and compute $\binom{N}{L}$ portfolios, and denote by $\mathcal{S} = (\mathcal{S}_1, \dots)$

Step 2: For each element of $\mathcal{S}$, find $\text{argmax}_{\widehat{\mathcal{S}}_i \in N \notin \mathcal{S}_i}$

Step 3: Iterate until reaching J, and select the portfolio with the highest utility.

When computing the optimal portfolio of size L out of N alternatives, the algorithm is initialized at $L \leq J$, where all portfolios of size L are computed and stored. Then, for each portfolio, we iteratively construct a size J portfolio using the MIA. The intuition for the algorithm stems from a direct characterization of the optimal portfolio. Setting $L = 2$, the optimal portfolio that consists of $(1, 2)$ is such that

[14]The analogy to the setting in Kelso and Crawford (1982) relates to the fact that interdependence between admission chances acts as an "externality" on the other components of an initial choice.

$$p_1 U_1 + (p_2 - p_1) U_2 > p_m U_m + (p_n - p_m) U_n \quad \forall m, n \in N \tag{9}$$

which implicitly defines a deviation term ϵ

$$\epsilon_{m,n} = p_1 U_1 + (p_2 - p_1) U_2 - p_m U_m - (p_n - p_m) U_n \quad \forall m, n \in N \tag{10}$$

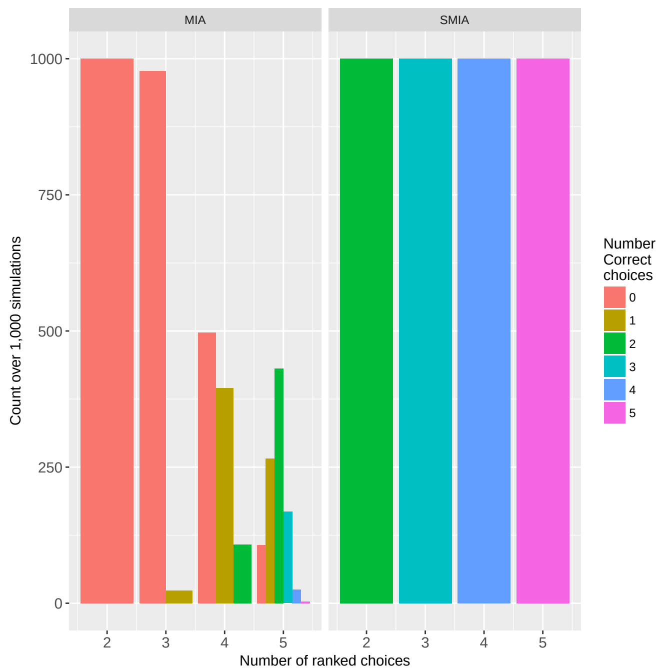
Moving forward, consider the optimal portfolio at step $L = 3$. Three potential motives drive the addition of a choice: expansion (first), diversification (interior), and back-up (last). For each of these motives, simple calculations define a range of deviation ϵ such that a portfolio that was not optimal in step $L - 1$ becomes optimal in step L. In other words, a nonoptimal portfolio in step $L - 1$ becomes optimal in step L only if it is close enough to the optimal solution in step $L - 1$. Because of this simple property, iterating on a large number of previously nonoptimal portfolios, yields the optimal portfolio. In practice, we may not need to iterate on all portfolios, but the optimal number depends on the level of dispersion in school utility and admission chances. In contrast to the true problem, where the number of operations increases at the combination rate, our algorithm requires an almost-constant number of operations regardless of the number of ranked choices, as the number of operations decreases with the number of ranked choices. As a consequence, we are able to cope with problems with a large number of alternatives. The most important parameter of the algorithm is the initial J, as it implies a trade-off between computation cost and efficiency. When J is equal to $N - 1$, the algorithm coincides with the optimal solution but requires as many operations as the standard brute force approach. On the contrary, when the difference between J and N increases the algorithm is likely to fail to recover the optimal solution.

Compared to MIA under correlated admission chances, our method yields portfolios which are substantially more selective. Specifically, once the motive of diversification is satisfied by a portfolio, average utility choices should be progressively replaced by higher utility choices that come with lower admission chances. MIA fails to accomodate for the latter as there is no restoring force to consider better alternatives that were not considered before. This is well illustrated by our running example.

4.2.3 Monte Carlo Simulations

Since the SMIA solution method is based on an approximation, we ascertain its performance using Monte Carlo simulations. Because the exact solution can be computed only when the number of available options is relatively small, we set $N = 100$. Individuals' utilities are drawn from a normal distribution with mean 10, and standard deviation 3, and admission chances are symmetrically obtained.

Figure 2: Monte Carlo Performance



Simulation based on 100 available options. The algorithm is initialized at $J = 2$.

Figure 2 reports the performance of the SMIA and MIA against the exact solution. SMIA performs perfectly in all instances with 100 available options, while MIA fails. Based on our experience with SMIA, the method has three important properties: (i) performance increases as the distance between the initialization step and the number of ranked choices decreases,[15] (ii) performance is not affected as the number of available options increases, and (iii) computation time increases with the number of alternatives but only at an additive rate.

[15]In our Monte Carlo simulations, the algorithm becomes unstable and sensitive to the initialization step once the number of ranked choices exceeds 10.

5 Estimation

This section describes our Simulated Methods of Moments (SMM) estimation method. The goal is to estimate the preferences and admission chances parameters given the optimal portfolio generation process for exogenous state variables (school capacity and individual observed heterogeneity). The next subsection describes the SMM technique in more detail. The following subsections describe the construction of the moments and discuss identification. For inference, we use a parametric bootstrap using 500 replications.

5.1 Simulated Method of Moments

We estimate the model by Simulated Method of Moments.[16] That is, we match the empirical characteristics of student ranked choices to their theoretical counterparts generated by the model. Formally, let θ denote the set of parameters to be estimated, the criterion function is given by:

$$\mathcal{L}(\theta) = -\frac{1}{2}\big(\widehat{m} - m(\theta)\big)^T \widehat{W}^{-1}\big(\widehat{m} - m(\theta)\big) \tag{11}$$

where $\widehat{m}$ is a set of empirical moments, $\widehat{W}$ is the weighting matrix, and $m(\theta)$ is the average of moments constructed from a large number of simulated portfolios. Specifically, for each draw of the error term, and a guess of preference parameters, we construct a portfolio of schools, generate the corresponding moments, and repeat for a large number of error draws.

5.2 Moments

The vector of empirical moments $\widehat{m}$ consists of empirical analogs of three sets of moments, calculated separately across individuals and then averaged across ordered ranked choices. For any simulated portfolio, $\mathcal{S} = \{\mathcal{S}_n\}_1^6$, the set of moments is given by:

[16] In a previous version, we implemented a Minimum Distance Estimator. However, this criterion function defined as a sum of 0 and 1 turned out to be highly convex, making the minimization extremely challenging.

1. Moments of the distribution of schools' observable characteristics

$$\widehat{m}_n = \frac{1}{I}\sum_i \left(Z_{ij}(\mathcal{S}_{i,n})\right) \tag{12}$$

2. Within ranked choices variance of schools' observable characteristics

$$\widehat{V}_n = \frac{1}{I}\sum_i \left(Z_{ij}(\mathcal{S}_{i,n}) - \overline{Z}_j\right)^2 \tag{13}$$

3. Moments of the joint distribution of students' and schools' observable characteristics

$$\widehat{J}_n = \frac{1}{I}\sum_i X_i Z_{ij}(\mathcal{S}_{i,n}) \tag{14}$$

In our final application, we include a total of 6 school characteristics (quality, boarding, old, general arts, general science, technical), and interactions between the first 3 of these school characteristics and 4 individual characteristics (male, testscore, JHS score, JHS public), which yields a total of 21 parameters (including the three parameters that characterize admission chances), and numerous potential moments. To reduce the number of moments, but also to validate the performance of our moments, we use only moments from the first, third, and last choices. Moments from the second, fourth and fifth choices are used for external validity analysis.

5.3 Identification

We consider the identification problem of our parameters of interest. The main challenge consists of determining whether a student's choice is driven by tastes over the attributes of a school or perceived admission chances. In this section, we show how the theoretical properties of the model can be used to identify these parameters separately. The proof is constructive, we demonstrate that preferences and admission parameters enter non-linearly in the value of a portfolio, and in turn show that we can sequentially pin down these parameters for identification purposes.

To develop an intuition for the identification, and to ease exposition, consider a portfolio that consists of two elements.[17] The utility individual i derives from choosing a school j is given by

$$\begin{aligned} U_{ij} &= \overline{U}_{ij} + \epsilon_{ij} \\ &= W_{ij}'\beta + \epsilon_{ij}. \end{aligned}$$

where $W_{ij} = (Z_j, d(l_i, l_j), Z_j X_i)$ and $\beta_{[d(l_i,l_j)]} = -1$. The utility individual i derives from submitting a portfolio $\mathcal{S}$ consisting of schools (j, k):

$$V_{i\mathcal{S}} = P(\psi_i - \Psi_j > \eta_{ij}) \left(\overline{U}_{ij} + \epsilon_{ij}\right) + \left[P(\psi_i - \Psi_k > \eta_{ik}) - P(\psi_i - \Psi_j > \eta_{ij})\right] \left(\overline{U}_{ik} + \epsilon_{ik}\right)$$

Denoting by σ, the parameters that govern uncertainty

$$V_{i\mathcal{S}} = p_{ij}(\sigma) \left(\overline{U}_{ij} + \epsilon_{ij}\right) + \left[p_{ik}(\sigma) - p_{ij}(\sigma)\right] \left(\overline{U}_{ik} + \epsilon_{ik}\right)$$

After a few steps of algebra, we obtain

$$V_{i\mathcal{S}} = \overline{W}_{i\mathcal{S}}(\sigma)'\beta + \bar{\epsilon}_{i\mathcal{S}}(\sigma) \tag{15}$$

where

$$\begin{aligned} \overline{W}_{i\mathcal{S}}(\sigma) &= p_{ij}(\sigma) W_{ij} + \left[p_{ik}(\sigma) - p_{ij}(\sigma)\right] W_{ik} \\ \bar{\epsilon}_{i\mathcal{S}}(\sigma) &= p_{ij}(\sigma)\epsilon_{ij} + \left[p_{ik}(\sigma) - p_{ij}(\sigma)\right] \epsilon_{ik} \\ p_{ij}(\sigma) &= 1 - F_\eta \left(\frac{\psi_i - \Psi_j}{\exp(\sigma_0 + \sigma_1\psi_i + \sigma_2\psi_m)}\right). \end{aligned}$$

Equation 15 shows that the value of a portfolio is a nonlinear function of elements that depends on admission chances and preferences parameters. Importantly, preferences and admission chances are linked only through the school selection mechanism captured by $\overline{W}_{i\mathcal{S}}(\sigma)$. Note that preferences are identified up to a scale as the model with $\tilde{\beta} = a\beta$ and $\tilde{\epsilon}_{ij} = a\epsilon_{ij}$ yields the same choice. The normalization of the coefficient on distance to "-1" serves that purpose and utilities are measured in terms of willingness to travel.

Now, let's consider an agent's decision to select one portfolio $\mathcal{S}$ over an alternative portfolio $\mathcal{S}' = (j', k')$. $\mathcal{S}$ is preferred over $\mathcal{S}'$ if and only if $V_{i\mathcal{S}} > V_{i\mathcal{S}'}$.

[17]The proof can be generalized to the case with n elements at no cost.

Hence the probability that agent i chooses the portfolio $\mathcal{S}$ is given by

$$\Pr\left(\left[\overline{W}_{i\mathcal{S}}(\sigma)-\overline{W}_{i\mathcal{S}'}(\sigma)\right]\beta+\overline{\epsilon}_{i\mathcal{S}}(\sigma)-\overline{\epsilon}_{i\mathcal{S}'}(\sigma)>0,\forall\mathcal{S}'\neq\mathcal{S}\right)$$

Writing the variance-covariance matrix further confirms that admission probabilities are independent of preference parameters.[18] Thus, identification of the model can be obtained using two layers of variation in the data. First, admission-probability parameters are identified from variation in school portfolios within groups of students with similar characteristics. Second, preference parameters are identified from variation in school portfolios across students with different observed attributes. We develop these arguments more explicitly in the following.

The observed component of the utility from choosing a specific portfolio is given by $\left[\overline{W}_{i\mathcal{S}}(\sigma)-\overline{W}_{i\mathcal{S}'}(\sigma)\right]\beta$, and since unobserved shocks, which enter non-linearly, depend only on admission-probability parameters σ, it follows that variation in school portfolios across similar students would indeed identify σ. More specifically, the constant parameter σ_0 is identified from variation in school portfolios across students with similar characteristics; σ_1 from the variation in school portfolios of students with the same set of observed attributes except individual test scores, σ_2 is identified from variation across students with similar observed characteristics but who attended different middle schools.

Given admission-probability parameters σ are identified, the decision of agents will be based on $\left[\overline{W}_{i\mathcal{S}}(\sigma)-\overline{W}_{i\mathcal{S}'}(\sigma)\right]\beta$. The identification of preference parameters follows from the principle of revealed preferences. Cross-sectional heterogeneity in the mapping between individual attributes, school characteristics and portfolios composition will identify preference parameters β. More precisely, the

[18] Under the assumption that $\epsilon_{ij}\overset{\text{i.i.d.}}{\sim}N(0,1)$ and for two alternative portfolios $\mathcal{S}=(j,k)$ and $\mathcal{S}'=(j',k')$. The elements in the variance-covariance matrix of choice-specific unobserved shocks $\overline{\epsilon}_i(\sigma)=\left[\overline{\epsilon}_{i\mathcal{S}}(\sigma)\right]_{\mathcal{S}}$ are given by:

$$\text{Cov}(\overline{\epsilon}_{i\mathcal{S}}(\sigma),\overline{\epsilon}_{i\mathcal{S}'}(\sigma))=\begin{cases} p_{ij}(\sigma)+\left(p_{ik}(\sigma)-p_{ij}(\sigma)\right)^2 & \text{if } \mathcal{S}=\mathcal{S}' \\ p_{ij}(\sigma) & \text{if } j'=j \text{ and } \mathcal{S}\cap\mathcal{S}'=\{j\} \\ \left[p_{ik}(\sigma)-p_{ij}(\sigma)\right]\left[p_{ik}(\sigma)-p_{ij'}(\sigma)\right] & \text{if } k'=k \text{ and } \mathcal{S}\cap\mathcal{S}'=\{k\} \\ p_{ij}(\sigma)\left[p_{ij}(\sigma)-p_{ij'}(\sigma)\right] & \text{if } k'=j \text{ and } \mathcal{S}\cap\mathcal{S}'=\{j\} \\ p_{ik}(\sigma)\left[p_{ik'}(\sigma)-p_{ik}(\sigma)\right] & \text{if } k=j' \text{ and } \mathcal{S}\cap\mathcal{S}'=\{k\} \\ 0 & \text{if } \mathcal{S}\cap\mathcal{S}'=\emptyset \end{cases}$$

set of preference parameters for school attributes is identified with the variation of school portfolios across all the students, while interaction parameters, $\{\Gamma_k\}_k$, are identified from variation in the hyper-plane formed by student and school attributes.

6 Results

The presentation of results is as follows. In the first part, we report our parameter estimates and the second part analyzes the capacity of our model to fit the data. We focus on three specifications that correspond to the different solution methods: SMIA, MIA, and the nonstrategic (truth telling) model. For estimation, we use the same set of moments allowing us to compare the different models.

6.1 Parameters

The three columns in Table 5 present parameter estimates respectively for SMIA, MIA, and the nonstrategic model. Although slightly better for SMIA, the moment criteria for SMIA and MIA are relatively close, 0.103 vs 0.115, indicating that both solution methods might fit individual behavior well. Contrastingly, the moment criterion for the nonstrategic model is 7 times higher.

We start by discussing the parameters from the SMIA model in column 1. The first three parameters characterize the standard deviation of the error term in students' beliefs about their admission chances. Consistent with qualitative evidence, our estimates suggest that students with higher test scores and those from higher quality junior high schools have less uncertainty about their admission chances. The remaining parameters characterize preferences for school attributes. On average, students prefer higher quality schools, boarding schools, and older schools established before Ghana gained independence. They have a significant preference for general arts and a strong negative taste for technical programs. Compared to these average preferences, students with higher test scores, those from higher-performing junior high schools, and those in public schools place relatively more value on school quality and boarding facilities, but have a weaker preference for older schools. Finally, male students have a stronger preference for school quality and older schools relative to females, with a significantly weaker preference for boarding schools.

For comparison, column 2 reports estimated parameters under the assumption that students select their portfolios using the MIA approach. In contrast to our SMIA estimates, the associated MIA parameters imply that students with higher test scores face more uncertainty than students with lower test scores. The latter finding may be explained by the fact that higher ability students are more likely to diversify their portfolios under interdependent admission chances, and hence induce more variance in portfolio quality, which may be interpreted as additional uncertainty in a model that doesn't account for this correlation in admission chances. Additionally, the interaction terms with school characteristics suggest that higher performing students place significantly less weight on school quality and boarding facilities, but have a stronger preference for attending older schools.

Finally, we consider a behavioral model in which students are completely truthful and simply list the six programs that would give them the highest utility. In this model, students do not consider their admission chances but instead ignore the possibility of being unassigned after exhausting their six choices. Under this specification, several parameters turn out be relatively different in sign and magnitude to estimates obtained SMIA, and MIA. Additionally, the moment criterion is extremely high pointing to a poor fit to the data. In particular, the parameters estimating heterogeneity by JHS background differ most strikingly.

6.2 Model Fit

To further validate our use of SMIA as a solution method, Appendix Table A.1 presents evidence on how well our model fits the data. We use our model to generate the distribution of attributes for the first, third, and sixth ranked choice in each portfolio and compare these with the real data. Our model fits the distribution of distance for each chosen school reasonably well. Both the data and model suggest that the most and least preferred schools are the farthest away, with the mid-ranked choices being the closest. The model captures the sharp decline in school quality over ranked choices. Nonetheless, the fit for average quality of selected choices is overstated, more seriously for lower ranked choices. SMIA thus generates portfolios that are substantially more selective than observed choices.

Table 5: Estimation Parameters

	Solution Method		
	SMIA	MIA	Non-strategic
Admission chances parameters			
Constant	-1.579^{***}	-2.600^{***}	
	(0.003)	(0.011)	
test score	-0.217^{***}	0.392^{***}	
	(0.049)	(0.008)	
junior school quality	-1.559^{***}	-1.375^{***}	
	(0.140)	(0.006)	
Preferences parameters			
quality	0.132^{***}	0.275^{***}	1.681^{***}
	(0.016)	(0.011)	(0.413)
boarding	0.559^{***}	0.323^{***}	0.221^{***}
	(0.021)	(0.009)	(0.047)
old	0.617^{***}	-0.196^{***}	0.402^{***}
	(0.051)	(0.010)	(0.024)
general arts	0.066^{***}	0.029^{***}	0.010
	(0.004)	(0.006)	(0.023)
general science	-0.018	-0.009	-0.043
	(0.040)	(0.012)	(0.044)
technical	-0.655^{***}	-0.613^{***}	-0.461
	(0.066)	(0.010)	(0.361)
quality×testscore	1.152^{***}	-1.157^{***}	1.012^{***}
	(0.024)	(0.008)	(0.054)
boarding×testscore	0.943^{***}	-0.295^{***}	0.029
	(0.051)	(0.014)	(0.018)
old×testscore	-1.045^{***}	1.887^{***}	-0.754^{***}
	(0.185)	(0.074)	(0.027)
quality×male	0.930^{***}	0.579^{***}	-0.290^{***}
	(0.014)	(0.012)	(0.037)
boarding×male	-0.345^{***}	-0.232^{***}	0.751^{***}
	(0.039)	(0.006)	(0.164)
old×male	1.217^{***}	-0.474^{***}	-0.872^{***}
	(0.193)	(0.074)	(0.049)
quality×juniorschoolscore	0.628^{***}	-0.239^{***}	-1.047^{***}
	(0.057)	(0.010)	(0.383)
boarding×juniorschoolscore	1.015^{***}	0.060^{***}	-0.360^{***}
	(0.052)	(0.013)	(0.029)
old×juniorschoolscore	-1.804^{***}	-0.525^{***}	0.571^{***}
	(0.139)	(0.011)	(0.012)
quality×juniorschoolpublic	1.509^{***}	0.437^{***}	-1.492^{***}
	(0.026)	(0.091)	(0.452)
boarding×juniorschoolpublic	0.580^{***}	0.024	-0.487^{***}
	(0.073)	(0.099)	(0.043)
old×juniorschoolpublic	-1.655^{***}	-0.343^{***}	-0.467^{***}
	(0.141)	(0.012)	(0.080)
Moment criteria	0.103	0.115	0.697
Note:			*p<0.1; **p<0.05; ***p<0.01

Beyond distance and quality, the actual and predicted school profiles (namely availability of boarding facilities and age) are quite close. We also predict the academic program choice patterns moderately well. General arts is substantially more popular than general science in both cases, and technical programs attract very little demand. Panel B of Appendix Table A.1 presents additional evidence on out-of-sample fit, using data on non-targeted moments (characteristics of the second, fourth, and fifth ranked choices). The main patterns in the targeted moments largely hold with the non-targeted moments.

By contrast, Appendix Table A.2 presents evidence on MIA. Although the fit for targeted moments is rather good, the out of sample fit is noticeably poor. For example, the pattern of declining school quality does not hold once all six choices are considered. Appendix Table A.3 reports the fit of the nonstrategic model. The nonstrategic model underpredicts the distance between students' junior high schools and their selected senior high school choices. More importantly, the nonstrategic model either does not generate enough variation (quality), or predicts too much variation (distance). These patterns are explained by the magnitude of individual taste for school quality, which can make up to 90% of some individuals' utilities. Additionally, this model does a poor job of matching the relatively high academic quality of students' first choice schools versus their lowest ranked choices.

Ultimately, we find that our optimal portfolio choice model better approximates actual school choice behavior. As expected, the nonstrategic model appears better at capturing extreme behaviors such as individuals applying only to the best schools.

7 Policy Experiments

Having examined the basic implications of our model, we turn to the task of simulating the welfare effects of two policy experiments – relaxing constraints on the number of choices students can list and reducing uncertainty about admission chances. To assess the welfare impacts of each policy, we predict student choices in a given policy environment and then simulate the assignment process to determine students' admission outcomes. We then calculate the utility each student would derive from their assigned school and estimate total welfare under each regime. In each case, we assume that the preference parameters estimated from the sophisticated model (SMIA) reflect students' true underlying references.

This section proceeds as follows. We first outline the assignment mechanism and specify our welfare function. We then discuss our simulation of the effects of increasing the number of choices and reducing uncertainty.

7.1 Assignment Mechanism and Welfare

Ghana's school choice system uses a student-proposing deferred acceptance algorithm to assign students to schools in the spirit of the matching procedure derived by Gale and Shapley (1962). The algorithm proceeds as follows:

- Step 1: Each student i applies to the first school in her ordered portfolio of choices. Each school s tentatively assigns its seats to applicants one at a time in order of students' exam scores, and rejects any remaining applicants once all of its seats are tentatively assigned.
- Step k: Each student who was rejected in round $k-1$ applies to the next school in her ordered portfolio of choices. Each school compares the set of students it has been holding to the set of new applicants. It tentatively assigns its seats to these students one at a time in order of students' exam scores and rejects remaining applicants once all of its seats are tentatively assigned.
- The algorithm terminates when no spaces remain in any of the choices selected by rejected students. Each student is then assigned to her final tentative assignment.

In practice, students who are unassigned at the end of the algorithm are administratively assigned to a program with remaining vacancies. In our simulations, we assign these students the outside option.

We propose a simple utilitarian welfare function $\mathcal{W}$ that aggregates individual utilities:

$$\mathcal{W} = \sum_i U_i^* \tag{16}$$

where U_i^* is the utility individual i derives from the school he was assigned. Total welfare is therefore the sum of individual utilities.

7.2 Expanding the Number of Listed Choices

To evaluate the effects of constraints on choices, we increase the number of choices students can submit incrementally from 1 to 8 and then allow students to submit an unlimited number of choices. This counterfactual exercise examines an aspect of mechanism design that policymakers often manipulate in practice. For example, students in Ghana were allowed to submit 6 senior high school choices in 2008, but the permitted length of ranked lists has fluctuated between 3 and 6 choices since 2005. Pathak and Sonmez (2013) document several similar policy changes in their review of school admissions reforms.

The effect of expanding the number of listed choices is theoretically ambiguous since the policy change will produce losers and winners on both the intensive and extensive margins. A student previously administratively assigned can get admission to a ranked option (extensive gain), while a student previously assigned can get administratively assigned (extensive loss). Similarly, a student can get admitted to a better school (intensive gain) while another one might get admitted into an inferior option (intensive loss). The final effect depends on the respective strength of each of these components.

Table 6: Constraints and Welfare

	Choices								
	1	2	3	4	5	6	7	8	J
Unmatched	0.72	0.55	0.45	0.37	0.31	0.15	0.13	0.09	0.00
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.05
1st Qu.	0.90	0.96	0.99	1.00	1.00	1.19	2.05	2.29	3.11
Median	1.00	1.00	2.01	2.47	2.62	2.74	2.86	2.92	3.45
Mean	1.43	1.81	2.08	2.29	2.45	2.58	2.72	2.82	3.49
Sd	1.07	1.24	1.31	1.31	1.30	1.26	1.20	1.15	0.73
3rd Qu.	1.76	2.74	3.01	3.19	3.32	3.39	3.45	3.51	3.86
Max	6.59	6.59	6.59	6.59	6.59	6.59	6.59	6.59	6.59
Total	182675	230964	266277	292125	312678	330223	347589	361073	445484

Note: The first row of this table reports the share of unassigned students. Rows 2-8 report the distribution of student welfare. The final row reports total student welfare. Column J indicates unconstrained choices.

Table 6 displays the change in the distribution of utilities under each regime. Expanding the number of choices from 1 to 2 reduces the share of unassigned students from 72 to 55 percent and increases total welfare by 126 percent. Although total welfare is a concave function of the number of choices, there are still significant gains from allowing students to submit an unrestricted number of choices. All students get assigned to a school and total welfare is almost 2.5 times as high as with only one choice.

Table 7: Constraints and Redistribution

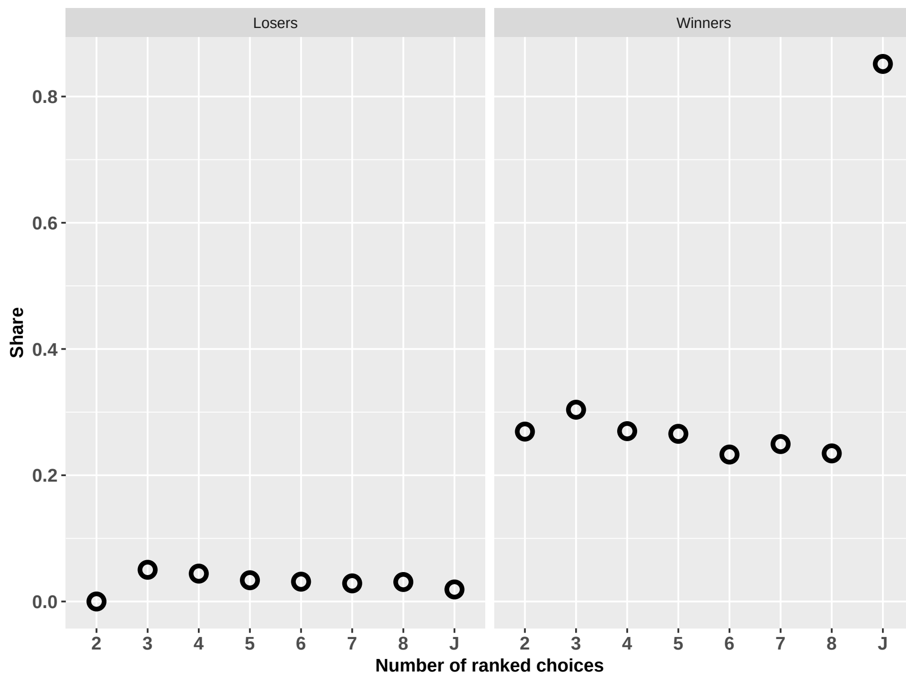
	Choices								
	1	2	3	4	5	6	7	8	J
0-10	3.1	5.2	4.2	4.4	4.7	4.9	5.7	5.8	8.7
10-20	3.1	4.4	4.9	5.3	5.8	6.3	6.2	6.4	8.9
20-30	4.2	5.2	6.5	6.4	6.7	6.3	6.2	6.1	8.8
30-40	3.1	6.2	6.3	5.6	5.9	6.4	6.5	6.4	9.7
40-50	4.0	5.8	5.8	6.4	6.1	6.1	7.0	8.8	9.6
50-60	5.0	6.1	5.9	7.9	8.1	9.8	11.2	11.6	10.3
60-70	8.5	8.8	11.5	10.7	11.9	12.3	12.5	12.1	10.3
70-80	12.4	10.3	11.1	14.0	14.8	14.6	13.8	13.3	10.7
80-90	22.9	19.1	20.0	18.6	17.2	16.0	14.9	14.2	11.2
90-100	33.7	28.8	23.8	20.7	18.7	17.3	16.0	15.2	11.8

Note: Each row reports the share of total welfare accruing to students in the associated decile of the test score distribution.

Despite the overall increase in welfare, benefits are not distributed uniformly across individuals. In Table 7, we decompose welfare gain by student ability. Students in the bottom decile of the exam score distribution receive only 3 percent of total welfare when there is only 1 choice, while students from the top decile enjoy a third of total welfare. With unlimited choices, students from the bottom decile account for 9 percent of total welfare and students from the top decile enjoy 12 percent. Thus, allowing a larger number of choices predominantly benefits lower ability students and generates a more equal distribution of welfare.

Figure 3 separately reports welfare changes for winners and losers. We define losers as students who experience at least a 0.1 mile decline in utility and winners as those who experience at least a 0.1 mile gain. Less than 0.05 percent of students are losers, meanwhile a quarter of students benefit each time there is a marginal expansion in the number of permitted choices, and over 80 percent of students benefit from going from eight choices to an unrestricted number. Figure 4 further decomposes winners and losers based on changes on the extensive margin (going from being unassigned to assigned) and the intensive margin (gaining admission to a more preferred school, conditional on being assigned). Moving from 1 to 2 choices, most of the gains and losses result from changes in the extensive margin. In contrast, expanding from 3 choices onward, most students experience welfare changes due to intensive margin changes in the welfare derived from the school to which they are assigned.

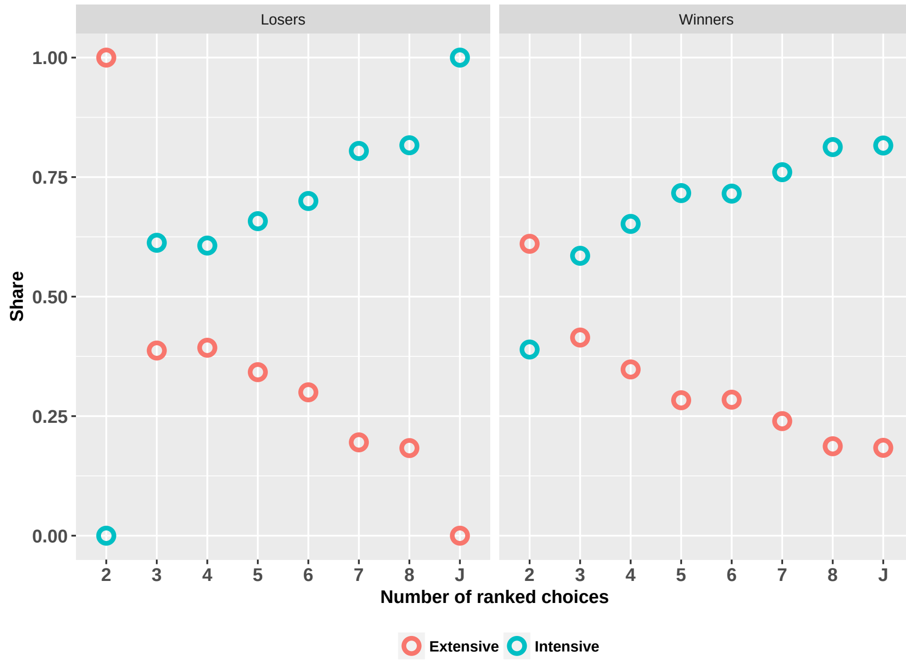
Figure 3: Winners and Losers



7.3 Turkish Experiment

Having analyzed the impact of constraints on the number of ranked choices, we turn now to consider an alternative policy option. Our model features a major source of inefficiency: uncertainty, which prompts agents to strategize over the submitted list of available choices. Since there is no simple strategy to select an optimal portfolio of schools, uncertainty implies potential strategic mistakes, that are welfare reducing. In this section, we propose an experiment to quantify the effect of information on individual choices and welfare.

Figure 4: Decomposition of intensive and extensive margins



As explained before, uncertainty in our setting originates from two distinct sources: individual uncertainty regarding realized test scores and aggregate uncertainty relating to the application behavior of other agents. While there exist obvious policy margins to deal with individual uncertainty, there is no easy way around aggregate uncertainty as participants in any large matching market are unlikely to have information on the preferences of all other agents. One could argue that uncertainty is irrelevant when the number of choices is unrestricted but policy makers have been reluctant to adopt such a policy and many school choice systems still restrict the number of ranked choices.[19] In this experiment, we effectively consider a compromise between the baseline case where uncertainty induces strategic mistakes and the alternative of unconstrained choices. We propose a realistic policy experiment, which eliminates individual uncertainty by informing students about their realized test scores (i.e., waiting for test score results to come out before students apply to schools). Then, we provide students with a subset of options that correspond to the programs where the prior year's admission cutoffs are lower than their realized test score. It turns out that university

[19] The few exceptions include school choice in Boston, Seattle, and Hungary.

admission in Turkey provides a real world example of this simulated mechanism (Akyol and Krishna, 2017).

Under binding constraints on the number of listed choices, agents may still be tempted to act strategically and not report true preferences. When these constraints are less severe, individual reporting behavior is less obvious, depending on the joint continuity of cutoffs and individual test scores. Specifically, under this mechanism strict truth telling is a symetric equilibrium, but not a dominant strategy, especially when the number of choices is small. As a consequence, we manipulate the length of the rank ordered list and let individuals report their true preferences, keeping in mind that truth telling is less likely to prevail when there is a smaller number of ranked choices.

Table 8: Uncertainty and Welfare

	Choices								
	1	2	3	4	5	6	7	8	J
Unmatched	0.50	0.27	0.18	0.12	0.07	0.04	0.02	0.01	0.00
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.05
1st Qu.	0.98	1.00	1.96	2.17	2.25	2.29	2.30	2.31	3.11
Median	1.05	2.56	2.67	2.70	2.72	2.73	2.73	2.73	3.45
Mean	1.97	2.43	2.62	2.72	2.79	2.84	2.86	2.88	3.49
Sd	1.29	1.23	1.14	1.06	0.98	0.92	0.89	0.86	0.73
3rd Qu.	2.93	3.27	3.35	3.36	3.36	3.36	3.36	3.36	3.86
Max	6.59	6.59	6.59	6.59	6.59	6.59	6.59	6.59	6.59
Total	252071	311102	335279	347663	356889	362870	366094	368302	445484

Note: The first row of this table reports the share of unassigned students. Rows 2-8 report the distribution of student welfare. The final row reports total student welfare. Column J indicates unconstrained choices.

As Table 8 indicates, half of students are assigned to a school even when restricted to listing only one choice. With eight choices, only 1 percent of students remain unassigned. This is a considerable improvement over the 72 percent and 18 percent respectively unassigned when students are choosing under uncertainty.

Analyzing the distributional consequences of this experiment, Table 9 reveals that lower ability students predominantly benefit compared to their higher ability peers – students in the bottom decile of test scores enjoy 6 percent of total welfare when students can submit only once choice, while students in the top decile enjoy 17.7 percent of total welfare. Although higher performing students maintain

Table 9: Uncertainty and Redistribution

	Choices								
	1	2	3	4	5	6	7	8	J
0-10	6.0	5.7	6.0	6.5	6.7	6.8	6.8	6.9	8.7
10-20	5.8	5.7	5.7	5.7	6.3	6.8	7.3	7.6	8.9
20-30	6.4	6.7	6.8	7.2	7.5	7.7	7.8	7.8	8.8
30-40	7.8	8.7	9.0	9.1	9.1	9.1	9.1	9.1	9.7
40-50	9.4	10.5	9.8	9.8	9.6	9.4	9.4	9.3	9.6
50-60	8.9	7.8	9.5	10.0	10.4	10.6	10.5	10.4	10.3
60-70	11.9	12.7	11.8	11.4	11.1	10.9	10.8	10.8	10.3
70-80	11.5	12.0	12.2	12.1	11.9	11.7	11.6	11.5	10.7
80-90	14.5	14.0	13.8	13.4	13.0	12.8	12.7	12.6	11.2
90-100	17.7	16.3	15.4	14.9	14.5	14.2	14.1	14.0	11.8

Note: Each row reports the share of total welfare accruing to students in the associated decile of the test score distribution.

their advantage, the distribution is considerably more even than under the case of incomplete information. Expanding the number of choices students can list generally increases the welfare of students in the bottom half of the achievement distribution.

Figure 5 illustrates the distribution of winners and losers. Less than 10 percent of students lose with the expansion from 1 to 2 choices, and the share of losers steadily declines with each subsequent increase in the number of ranked choices. Given that total welfare is higher with only one ranked choice when students have reduced uncertainty compared to our baseline case, there are relatively fewer winners as the number of ranked choices increases. Nonetheless, a quarter of students benefit from the switch from 1 to 2 choices and there is a sizeable but declining pool of winners from further expansions. Once again, over 80 percent of students gain from unrestricting the number of ranked choices.

Turning to the decomposition of intensive and extensive margins in Figure 6, we observe losses occurring almost exclusively on the intensive margin while winners almost exclusively gain on the extensive margin. It is worth noting that intensive margin losses are relatively small. The lack of intensive margin gain is explained by inefficiencies within the current allocation system as some preferable schooling options may have remaining vacancies in equilibrium. In a final exercise, we attempt to introduce additional efficiency by expanding the set of

Figure 5: Distribution of welfare

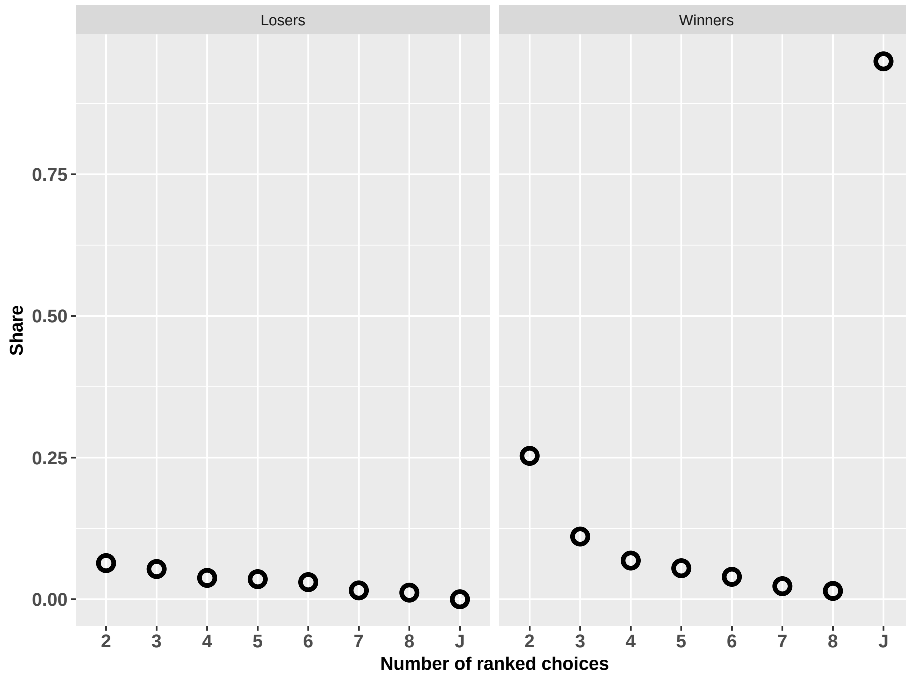
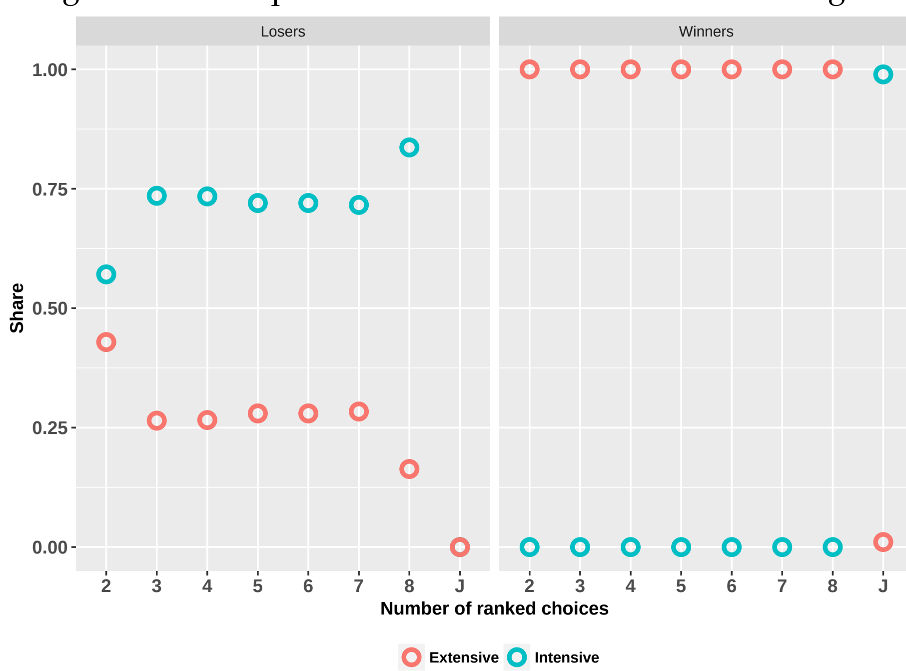


Figure 6: Decomposition of intensive and extensive margins



schools where individuals can apply.[20] In this exercise, we do not manipulate the number of ranked choices anymore, as the previous result shows that administrative assignment is less of a concern when students are allowed to submit 8 choices. Instead, we hold the number of permitted choices fixed at 8 and allow students to be increasingly ambitious in their application behavior.

Table 10: Uncertainty, Efficiency and Welfare

	Choices								
	Benchmark	5%	10%	15%	20%	25%	30%	35%	40%
Unmatched	0.010	0.006	0.002	0.000	0.000	0.004	0.025	0.054	0.090
Min	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1st Qu.	2.308	2.351	2.393	2.431	2.462	2.489	2.486	2.467	2.429
Median	2.730	2.779	2.833	2.895	2.962	3.021	3.042	3.047	3.038
Mean	2.881	2.937	2.990	3.043	3.091	3.127	3.103	3.056	2.988
Sd	0.860	0.866	0.869	0.872	0.877	0.886	0.924	0.978	1.043
3rd Qu.	3.362	3.418	3.478	3.539	3.594	3.646	3.649	3.636	3.609
Max	6.585	6.585	6.585	6.585	6.585	6.585	6.585	6.585	6.585
Total	368302	375365	382217	388933	395092	399672	396644	390622	381999

Note: Benchmark assumes students can submit 8 choices and list their most preferred alternatives from the set with prior admission cutoffs below their individual test score. In remaining columns, students chose from the set of alternatives with admission cutoffs below their test score $+X\%$.

Table 10 reports the welfare distribution and share of individuals administratively assigned when we allow students to apply to programs with a cutoff that is respectively 5%, 10%, 15%, . . . , 40% higher than their test score. Strikingly, the effect of allowing students to be increasingly ambitious is not monotonic. The share of administrative assignments falls and total welfare increases as students apply to programs with cutoffs up to 20% above their test score. Beyond this point however, administrative assignments begin to increase again and total welfare eventually starts to decline. Essentially, students benefit from being ambitious but not too ambitious – as the gap between a student's test score and program cutoffs becomes too large (in our simulation, at 25% and above), there is a cost to applying to programs definitively outside of reach. Table 10 also exhibits the trade-off faced by policymakers. That is, letting students apply to substantially

[20]We attempted to recover the "long-run" equilibrium of the matching procedure by updating cutoffs after each iteration. This procedure turns out to be extremely slow – after 10 iterations the share of administrative assignment was still around 20%, and total welfare is lower than the benchmark case.

more selective schools (40% for example) generates more total welfare than the benchmark case, despite the higher rate of administrative assignment.

8 Conclusion

In this paper, we propose a framework to estimate preferences and simulate the welfare effects of school choice policies in systems with costly choices and uncertainty about admission chances. We view students' ranking of schools as the solution to a portfolio choice problem and introduce a Submodular Improvement Algorithm to match theoretical predictions to empirical moments. Consequently, our estimated parameters are policy-invariant and can be used to ascertain the welfare implications of several policy experiments.

Our work addresses a longstanding challenge in the school choice literature. Since the first formalization of the school choice problem, both policymakers and matching theorists have considered incomplete information a daunting problem. Under incomplete information, a constrained number of choices leads to diversification of submitted choices. As such, empirical exercises that assume complete or imperfect information in a setting where students are uncertain about their admission chances may misconceive the utility individuals derive from schools, as illustrated by our estimated parameters.

We find that allowing students to rank an unlimited number of choices generates multiple benefits. Total welfare increases by a factor of 2.4 compared to a system where students can list only one choice, and the proportion of students who are unassigned falls from 72 percent to zero. This policy also reduces inequality as lower ability students experience larger improvements in welfare. Theoretically, expanding the number of choices could have ambiguous effects – improving match quality for students but eliminating an opportunity to signal preference intensity. We provide much needed empirical evidence on this open question. Further, we show that it is possible to design policies to accommodate uncertainty using simple and appealing strategies: providing information to students about their test scores and restricting students from applying to schools historically more selective than their test scores warrant yield substantial gains both on the intensive (quality of school) and extensive (probability of being assigned) margins.

The methods used in this paper provide several avenues for future research. Although our analysis focuses on two common features of coordinated school assignment systems, our empirical approach has broad applications. We estimate a model of optimal portfolio choice that allows us to conduct other experiments related to changes in mechanism design (such as restricting the categories of schools students can select from, using a lottery instead of merit-based admissions procedure, or modifying the assignment algorithm) as well as experiments related to institutional reforms (such as increasing the capacity of high-performing schools or constructing new schools in remote areas). We leave this for future work.

Our model rests on the assumption that agents understand the strategic implications of constraints on the number of ranked alternatives and act accordingly when constructing portfolios of schools. Because not all agents are likely to be sophisticated, it may be informative to estimate a model that includes heterogeneity in both preferences and sophistication. However, identification requires jointly observing data on preferences and beliefs about admission chances, which we unfortunately do not have in this context.

References

Abdulkadiroglu, A., Agarwal, N., and Pathak, P. (2015). The Welfare Effects of Coordinated School Assignment: Evidence from the NYC High School Match. NBER Working Papers 21046, National Bureau of Economic Research, Inc.

Abdulkadiroglu, A., Angrist, J., and Pathak, P. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, 82(1):137–196.

Abdulkadiroglu, A., Che, Y.-K., and Yasuda, Y. (2011). Resolving Conflicting Preferences in School Choice: The Boston Mechanism Reconsidered. *American Economic Review*, 101(1):399–410.

Abdulkadiroglu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match. *American Economic Review*, 99(5):1954–78.

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

Agarwal, N. and Somaini, P. (2014). Demand Analysis using Strategic Reports: An Application to a School Choice Mechanism. NBER Working Papers 20775, National Bureau of Economic Research, Inc.

Ajayi, K. F. (2013). School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana. Institute for Economic Development Discussion Paper 259, Boston University.

Akyol, P. and Krishna, K. (2017). Preferences, Selection, and Value Added: A Structural Approach. *European Economic Review*, 91:89–117.

Azevedo, E. M. and Leshno, J. D. (2016). A Supply and Demand Framework for Two-Sided Matching Markets. *Journal of Political Economy*, 124(5):1235–1268.

Berry, S. and Pakes, A. (2007). The Pure Characteristics Demand Model. *International Economic Review*, 48(4):1193–1225.

Calsamiglia, C., Fu, C., and Guell, M. (2014). Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives. Working Papers 2014-21, FEDEA.

Calsamiglia, C., Haeringer, G., and Klijn, F. (2010). Constrained School Choice: An Experimental Study. *American Economic Review*, 100(4):1860–74.

Chade, H., Lewis, G., and Smith, L. (2014). Student Portfolios and the College Admissions Problem. *Review of Economic Studies*, 81(3):971–1002.

Chade, H. and Smith, L. (2006). Simultaneous Search. *Econometrica*, 74(5):1293–1307.

Chen, Y. and Sonmez, T. (2006). School Choice: An Experimental Study. *Journal of Economic Theory*, 127(1):202–231.

Duflo, E., Dupas, P., and Kremer, M. (2017). The Impact of Free Secondary Education: Experimental Evidence from Ghana. Working paper, Stanford University.

Ergin, H. and Sonmez, T. (2006). Games of School Choice under the Boston Mechanism. *Journal of Public Economics*, 90(1-2):215–237.

Fack, G., Grenet, J., and He, Y. (2015). Beyond Truth-Telling: Preference Estimation with Centralized School Choice. CEPR Discussion Papers 10907, C.E.P.R. Discussion Papers.

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

He, Y. (2012). Gaming the Boston School Choice Mechanism in Beijing. TSE Working Papers 12-345, Toulouse School of Economics (TSE).

Kelso, A. S. J. and Crawford, V. P. (1982). Job Matching, Coalition Formation, and Gross Substitutes. *Econometrica*, 50(6):1483–1504.

Laporte, G. (1992). The Traveling Salesman Problem: An Overview of Exact and Approximate Algorithms. *European Journal of Operational Research*, 59(2):231–247.

Pallais, A. (2015). Small differences that matter: Mistakes in applying to college. *Journal of Labor Economics*, 33(2):493–520.

Pathak, P. A. and Shi, P. (2014). Demand Modeling, Forecasting, and Counterfactuals, Part I. NBER Working Papers 19859, National Bureau of Economic Research, Inc.

Pathak, P. A. and Sonmez, T. (2013). School Admissions Reform in Chicago and England: Comparing Mechanisms by Their Vulnerability to Manipulation. *American Economic Review*, 103(1):80–106.

Pop-Eleches, C. and Urquiola, M. (2013). Going to a Better School: Effects and Behavioral Responses. *American Economic Review*, 103(4):1289–1324.

Troyan, P. (2012). Comparing School Choice Mechanisms by Interim and Ex-Ante Welfare. *Games and Economic Behavior*, 75(2):936–947.

A Appendix

A.1 Details of MIA

The algorithm proceeds as follows.

Step 1: In the first step, we select the first ranked choice as the school that yields the maximum expected utility. Formally, we solve the following problem:

$$\mathcal{S}_1 = \underset{j}{\text{argmax}}\, q_j U_j$$

Then we store the first optimal portfolio as the school that yields the highest utility $\mathcal{U}_1 = (U_1^*) = U_j$, and the corresponding admission probability $\mathcal{Q}_1 = (q_1^*) = q_j$

Step 2: In step 2, the following problem is solved:

$$\mathcal{S}_2 = \underset{j \in J \notin \mathcal{S}_1}{\text{argmax}} \begin{cases} q_1^* U_1^* + \mathcal{P}(q_j | q_1^*) U_j & \text{if} \quad U_1^* \geq U_j \\ q_j U_j + \mathcal{P}(q_1^* | q_j) U_1^* & \text{if} \quad U_j > U_1^* \end{cases}$$

where $\mathcal{P}(q_j|q_1^*) = P(\hat{\psi} > \hat{\Psi}_j | \hat{\psi} < \hat{\Psi}_1)$ is the probability of being accepted in school j conditional on not being accepted in the first round choice. Given the outcome of this optimization problem, the optimal portfolio is updated $\mathcal{U}_2 = (U_1^*, U_j) = (U_1^*, U_2^*)$ with $U_1^* > U_2^*$. Note that the portfolio is arranged in descending order of utilities. The corresponding set of admission chances is given by: $\mathcal{Q}_2 = (q_1^*, q_2^*)$

Step 3: In the third step, we maximize over all schools that have not been selected in the previous two steps.

$$\mathcal{S}_3 = \underset{j \in J \notin \mathcal{S}_2}{\text{argmax}} \begin{cases} q_1^* U_1^* + \mathcal{P}(q_2^* | q_1^*) U_2^* + \mathcal{P}(q_j | q_2^*) U_j & \text{if} \quad U_2^* \geq U_j \\ q_1^* U_1^* + \mathcal{P}(q_j | q_1^*) U_j + \mathcal{P}(q_2^* | q_j) U_2^* & \text{if} \quad U_1^* > U_j > U_2^* \\ q_j U_j + \mathcal{P}(q_1^* | q_j) U_1^* + \mathcal{P}(q_2^* | q_1^*) U_2^* & \text{if} \quad U_j > U_1^* \end{cases}$$

Then store $\mathcal{U}_3 = (U_1^*, U_2^*, U_j) = (U_1^*, U_2^*, U_3^*)$ with $U_1^* > U_2^* > U_3^*$.

Step 4: In the fourth step, we select the school that yields the highest utility among all remaining schools.

$$\mathcal{S}_4 = \underset{j \in J \notin \mathcal{S}_3}{\operatorname{argmax}} \begin{cases} q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_j|q_3^*)U_j & \text{if} \quad U_3^* \geq U_j \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_j^*|q_2^*)U_j + \mathcal{P}(q_3^*|q_j)U_3^* & \text{if} \quad U_2^* > U_j > U_3^* \\ q_1^* U_1^* + \mathcal{P}(q_j|q_1^*)U_j + \mathcal{P}(q_2^*|q_j)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* & \text{if} \quad U_1^* > U_j > U_2^* \\ q_j U_j + \mathcal{P}(q_1^*|q_j)U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* & \text{if} \quad U_j > U_1^* \end{cases}$$

Then store $\mathcal{U}_4^* = (U_1^*, U_2^*, U_3^*, U_j) = (U_1^*, U_2^*, U_3^*, U_4^*)$ with $U_1^* > U_2^* > U_3^* > U_4^*$.

Step 5: The fifth step carries on using the same strategy as before.

$$\mathcal{S}_5 = \underset{j \in J \notin \mathcal{S}_4}{\operatorname{argmax}} \begin{cases} q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_j|q_4^*)U_j & \text{if} \quad U_4^* \geq U_j \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_j^*|q_3^*)U_j + \mathcal{P}(q_4^*|q_j)U_4^* & \text{if} \quad U_3^* > U_j > U_4^* \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_j^*|q_2^*)U_j + \mathcal{P}(q_3^*|q_j)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* & \text{if} \quad U_2^* > U_j > U_3^* \\ q_1^* U_1^* + \mathcal{P}(q_j|q_1^*)U_j + \mathcal{P}(q_2^*|q_j)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* & \text{if} \quad U_1^* > U_j > U_2^* \\ q_j U_j + \mathcal{P}(q_1^*|q_j)U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* & \text{if} \quad U_j > U_1^* \end{cases}$$

Then store $\mathcal{U}_5^* = (U_1^*, U_2^*, U_3^*, U_4^*, U_j) = (U_1^*, U_2^*, U_3^*, U_4^*, U_5^*)$ with $U_1^* > U_2^* > U_3^* > U_4^* > U_5^*$.

Step 6 : Finally, we maximize over the remaining schools to obtain a sixth choice.

$$\mathcal{S}_6 = \underset{j \in J \notin \mathcal{S}_5}{\operatorname{argmax}} \begin{cases} q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_5^*|q_4^*)U_5^* + \mathcal{P}(q_j|q_5^*)U_j & \text{if} \quad U_5^* \geq U_j \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_j|q_4^*)U_j + \mathcal{P}(q_5^*|q_j)U_5^* & \text{if} \quad U_4^* > U_j > U_5^* \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_j^*|q_3^*)U_j + \mathcal{P}(q_4^*|q_j)U_4^* + \mathcal{P}(q_5^*|q_4^*)U_5^* & \text{if} \quad U_3^* > U_j > U_4^* \\ q_1^* U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_j|q_2^*)U_j + \mathcal{P}(q_3^*|q_j)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_5^*|q_4^*)U_5^* & \text{if} \quad U_2^* > U_j > U_3^* \\ q_1^* U_1^* + \mathcal{P}(q_j|q_1^*)U_j + \mathcal{P}(q_2^*|q_j)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_5^*|q_4^*)U_5^* & \text{if} \quad U_1^* > U_j > U_2^* \\ q_j U_j + \mathcal{P}(q_1^*|q_j)U_1^* + \mathcal{P}(q_2^*|q_1^*)U_2^* + \mathcal{P}(q_3^*|q_2^*)U_3^* + \mathcal{P}(q_4^*|q_3^*)U_4^* + \mathcal{P}(q_5^*|q_4^*)U_5^* & \text{if} \quad U_j > U_1^* \end{cases}$$

Then store $\mathcal{U}_6 = (U_1^*, U_2^*, U_3^*, U_4^*, U_5^*, U_j) = (U_1^*, U_2^*, U_3^*, U_4^*, U_5^*, U_6^*)$ with $U_1^* > U_2^* > U_3^* > U_4^* > U_5^* >, U_6^*$.

A.2 Model Fit

Table A.1: Model Fit - SMIA

Panel A: Targeted moments

	Data			Model		
	choice 1	choice 3	choice 6	choice 1	choice 3	choice 6
$\widehat{m}$(distance)	0.079	0.07	0.071	0.12	0.099	0.21
$\widehat{V}$(distance)	0.11	0.098	0.063	0.15	0.093	0.19
$\widehat{m}$(quality)	0.62	0.5	0.35	0.8	0.65	0.52
$\widehat{V}$(quality)	0.2	0.18	0.11	0.1	0.14	0.17
$\widehat{m}$(boarding)	0.88	0.77	0.59	1	0.99	0.95
$\widehat{V}$(boarding)	0.32	0.42	0.49	0	0.11	0.21
$\widehat{m}$(old)	0.26	0.11	0.017	0.11	0.11	0.15
$\widehat{V}$(old)	0.44	0.31	0.13	0.31	0.31	0.36
$\widehat{m}$(general arts)	0.4	0.39	0.43	0.3	0.28	0.16
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.46	0.45	0.37
$\widehat{m}$(general science)	0.14	0.09	0.067	0.17	0.27	0.052
$\widehat{V}$(general science)	0.35	0.29	0.25	0.37	0.44	0.22
$\widehat{m}$(technical)	0.0082	0.01	0.00007	0	0	0
$\widehat{V}$(technical)	0.09	0.1	0.0084	0	0	0
$\widehat{J}$(quality)	1.3	1	0.74	1.7	1.4	1.1
$\widehat{J}$(boarding)	1.9	1.6	1.2	2.1	2.1	2
$\widehat{J}$(old)	0.57	0.22	0.037	0.21	0.21	0.3

Panel B: Non targeted moments

	Data			Model		
	choice 2	choice 4	choice 5	choice 2	choice 4	choice 5
$\widehat{m}$(distance)	0.075	0.061	0.069	0.11	0.1	0.11
$\widehat{V}$(distance)	0.1	0.094	0.063	0.13	0.093	0.11
$\widehat{m}$(quality)	0.55	0.44	0.37	0.71	0.59	0.53
$\widehat{V}$(quality)	0.19	0.17	0.11	0.13	0.15	0.16
$\widehat{m}$(boarding)	0.83	0.69	0.61	0.99	0.98	0.97
$\widehat{V}$(boarding)	0.38	0.46	0.49	0.078	0.14	0.17
$\widehat{m}$(old)	0.16	0.073	0.026	0.11	0.11	0.1
$\widehat{V}$(old)	0.36	0.26	0.16	0.31	0.31	0.3
$\widehat{m}$(general arts)	0.39	0.39	0.43	0.32	0.36	0.44
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.46	0.48	0.5
$\widehat{m}$(general science)	0.11	0.078	0.078	0.22	0.14	0.11
$\widehat{V}$(general science)	0.31	0.27	0.27	0.42	0.35	0.31
$\widehat{m}$(technical)	0.0088	0.015	0.000086	0	0	0
$\widehat{V}$(technical)	0.093	0.12	0.0093	0	0	0
$\widehat{J}$(quality)	1.1	0.93	0.78	1.5	1.3	1.1
$\widehat{J}$(boarding)	1.7	1.4	1.3	2.1	2	2
$\widehat{J}$(old)	0.33	0.15	0.059	0.21	0.2	0.19

Table A.2: Model Fit - MIA

Panel A: Targeted moments

	Data			Model		
	choice 1	choice 3	choice 6	choice 1	choice 3	choice 6
$\widehat{m}$(distance)	0.079	0.07	0.071	0.043	0.067	0.13
$\widehat{V}$(distance)	0.11	0.098	0.063	0.055	0.082	0.14
$\widehat{m}$(quality)	0.62	0.5	0.35	0.6	0.48	0.31
$\widehat{V}$(quality)	0.2	0.18	0.11	0.17	0.16	0.14
$\widehat{m}$(boarding)	0.88	0.77	0.59	0.89	0.78	0.55
$\widehat{V}$(boarding)	0.32	0.42	0.49	0.32	0.41	0.5
$\widehat{m}$(old)	0.26	0.11	0.017	0.27	0.12	0.036
$\widehat{V}$(old)	0.44	0.31	0.13	0.44	0.33	0.19
$\widehat{m}$(general arts)	0.4	0.39	0.43	0.54	0.41	0.24
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.5	0.49	0.43
$\widehat{m}$(general science)	0.14	0.09	0.067	0.095	0.1	0.063
$\widehat{V}$(general science)	0.35	0.29	0.25	0.29	0.3	0.24
$\widehat{m}$(technical)	0.0082	0.01	0.00007	0	0	0.021
$\widehat{V}$(technical)	0.09	0.1	0.0084	0	0	0.14
$\widehat{J}$(quality)	1.3	1	0.74	1.3	1	0.7
$\widehat{J}$(boarding)	1.9	1.6	1.2	1.8	1.6	1.1
$\widehat{J}$(old)	0.57	0.22	0.037	0.47	0.23	0.068

Panel B: Non targeted moments

	Data			Model		
	choice 2	choice 4	choice 5	choice 2	choice 4	choice 5
$\widehat{m}$(distance)	0.075	0.061	0.069	0.11	0.1	0.11
$\widehat{V}$(distance)	0.1	0.094	0.063	0.13	0.093	0.11
$\widehat{m}$(quality)	0.55	0.44	0.37	0.71	0.59	0.53
$\widehat{V}$(quality)	0.19	0.17	0.11	0.13	0.15	0.16
$\widehat{m}$(boarding)	0.83	0.69	0.61	0.99	0.98	0.97
$\widehat{V}$(boarding)	0.38	0.46	0.49	0.078	0.14	0.17
$\widehat{m}$(old)	0.16	0.073	0.026	0.11	0.11	0.1
$\widehat{V}$(old)	0.36	0.26	0.16	0.31	0.31	0.3
$\widehat{m}$(general arts)	0.39	0.39	0.43	0.32	0.36	0.44
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.46	0.48	0.5
$\widehat{m}$(general science)	0.11	0.078	0.078	0.22	0.14	0.11
$\widehat{V}$(general science)	0.31	0.27	0.27	0.42	0.35	0.31
$\widehat{m}$(technical)	0.0088	0.015	0.000086	0	0	0
$\widehat{V}$(technical)	0.093	0.12	0.0093	0	0	0
$\widehat{J}$(quality)	1.1	0.93	0.78	1.5	1.3	1.1
$\widehat{J}$(boarding)	1.7	1.4	1.3	2.1	2	2
$\widehat{J}$(old)	0.33	0.15	0.059	0.21	0.2	0.19

Table A.3: Model Fit - Nonstrategic

Panel A: Targeted moments

	Data			Model		
	choice 1	choice 3	choice 6	choice 1	choice 3	choice 6
$\widehat{m}$(distance)	0.079	0.07	0.071	0.026	0.031	0.038
$\widehat{V}$(distance)	0.11	0.098	0.063	0.052	0.056	0.058
$\widehat{m}$(quality)	0.62	0.5	0.35	0.51	0.52	0.51
$\widehat{V}$(quality)	0.2	0.18	0.11	0.31	0.27	0.26
$\widehat{m}$(boarding)	0.88	0.77	0.59	0.69	0.69	0.69
$\widehat{V}$(boarding)	0.32	0.42	0.49	0.46	0.46	0.46
$\widehat{m}$(old)	0.26	0.11	0.017	0.12	0.12	0.12
$\widehat{V}$(old)	0.44	0.31	0.13	0.32	0.32	0.32
$\widehat{m}$(general arts)	0.4	0.39	0.43	0.65	0.36	0.33
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.48	0.48	0.47
$\widehat{m}$(general science)	0.14	0.09	0.067	0.14	0.079	0.073
$\widehat{V}$(general science)	0.35	0.29	0.25	0.35	0.27	0.26
$\widehat{m}$(technical)	0.0082	0.01	0.00007	0	0	0
$\widehat{V}$(technical)	0.09	0.1	0.0084	0	0	0
$\widehat{J}$(quality)	1.3	1	0.74	0.98	1	1
$\widehat{J}$(boarding)	1.9	1.6	1.2	1.6	1.6	1.6
$\widehat{J}$(old)	0.57	0.22	0.037	0.12	0.12	0.12

Panel B: Non targeted moments

	Data			Model		
	choice 2	choice 4	choice 5	choice 2	choice 4	choice 5
$\widehat{m}$(distance)	0.075	0.061	0.069	0.11	0.1	0.11
$\widehat{V}$(distance)	0.1	0.094	0.063	0.13	0.093	0.11
$\widehat{m}$(quality)	0.55	0.44	0.37	0.71	0.59	0.53
$\widehat{V}$(quality)	0.19	0.17	0.11	0.13	0.15	0.16
$\widehat{m}$(boarding)	0.83	0.69	0.61	0.99	0.98	0.97
$\widehat{V}$(boarding)	0.38	0.46	0.49	0.078	0.14	0.17
$\widehat{m}$(old)	0.16	0.073	0.026	0.11	0.11	0.1
$\widehat{V}$(old)	0.36	0.26	0.16	0.31	0.31	0.3
$\widehat{m}$(general arts)	0.39	0.39	0.43	0.32	0.36	0.44
$\widehat{V}$(general arts)	0.49	0.49	0.49	0.46	0.48	0.5
$\widehat{m}$(general science)	0.11	0.078	0.078	0.22	0.14	0.11
$\widehat{V}$(general science)	0.31	0.27	0.27	0.42	0.35	0.31
$\widehat{m}$(technical)	0.0088	0.015	0.000086	0	0	0
$\widehat{V}$(technical)	0.093	0.12	0.0093	0	0	0
$\widehat{J}$(quality)	1.1	0.93	0.78	1.5	1.3	1.1
$\widehat{J}$(boarding)	1.7	1.4	1.3	2.1	2	2
$\widehat{J}$(old)	0.33	0.15	0.059	0.21	0.2	0.19