School Choice Under Imperfect Information.*

Kehinde Ajayi† Modibo Sidibé‡

December 16, 2019

Abstract

As in many school districts around the world, prospective high-school students in Ghana are assigned to schools through a centralized system. Using administrative data on applications, we report that virtually all students adopt a weakly dominated strategy, and matching outcomes show that approximately 15% of students end up unassigned, while almost half of schools have at least 1 vacancy. In order to rationalize choices in this setting, we build and estimate a model, where students engage in a costly search process to acquire information over school characteristics. The key insight of the model is that schooling decisions are exerted without the full examination of all available options, which may lead to sub-optimal choices. Our empirical application documents a substantial welfare loss: distance traveled to schools could be divided by 4. Counterfactual simulations show that if a planner were to restrict choices and assign the highest test score student to the most selective school, welfare would increase by 72%.

Keywords: school choice, uncertainty, consideration set, search.

JEL Classification Number: C53, D61, I20.

*This paper has benefited from helpful discussions with Attila Ambrus, Pat Bayer, Jacob Leshno, Adam Rosen and Ran Shorrer as well as seminar and conference participants at Duke University, University of Virginia, University of Wisconsin, UCLA and PeGnet.

†The World Bank, 1818 H Street NW, Washington, DC 20433. Email: kajayi@worldbank.org.

‡Department of Economics, Duke University, 213 Social Sciences Building, 419 Chapel Drive, Box 90097, Durham, NC 27708. Email: modibo.sidibe@duke.edu.

Introduction

Over the last 30 years, choice has become a key aspect of the assignment of students to schools. As such, most public school systems around the world are organized through a centralized coordinated assignment mechanism. Recent empirical research on school choice highlights the potential welfare gain.[1]

In 2005, Ghana introduced the national Computerized School Selection and Placement System (CSSPS) to match prospective students to high schools. The stated goal was to increase equity and access to quality senior high schools. The matching is based on the serial dictatorship algorithm, where priorities are determined by the student score on the Basic Education Certification Examination (BECE). Every spring, several hundred thousands of students submit a wish list of schools, and gain admission into one school at the end of summer, making it *de facto* one of the largest matching systems in the world.

Yet, throughout the process, logistical considerations outweigh efficiency concerns. First, the timing introduces uncertainty, as rank-order lists (ROLs) are submitted prior to the examination that determined priority scores. Then, constraints were imposed in the length of rank ordered lists (3 in 2005, 4 in 2007, then 6 in 2008), which prompt agents to strategize over their submitted list. The short history of the program combined with the potential low involvement of parents may have worsened the potential welfare consequences of these implementation issues.

Three years after the introduction of the CSSPS, an analysis of application lists demonstrates that established principles of students application under constraints such as the ordering of choices in admission chances are violated for almost 95% of the students, and as much as half (50%) of the students could have been admitted into a more selective school by changing the ordering of schools in their application. Matching outcomes show that approximately 15% of students end up unassigned.[2] The application behavior of students led to churning across admission cutoffs, especially for low and medium selectivity schools, and almost 50% of schools end up with at least 1 vacancy (including the very best schools).

As many school districts around the world engage in policy reforms that include centralized allocation of students to schools, uncertainty over admission outcomes and applications mistakes (Pallais, 2015; Lucas and Mbiti, 2012) are likely to affect the welfare of students. While there is no consensus on the causes of these inefficiencies, informational constraints (Hastings and Weinstein, 2008; Hoxby and Turner, 2015) pose an important challenge, especially for students from disadvantaged backgrounds and may negate the

[1] see Agarwal and Somaini (2019); Abdulkadiroglu et al. (2011) for reviews.

[2] 30% of students were administratively assigned under the decentralized application in NYC, motivating the switch to a coordinated assignment according to Abdulkadiroglu et al. (2005).

initial objectives of school choice reforms. Yet, the school choice literature is cast in a setting where students and parents are informed about all the schooling opportunities.[3] In this paper, we introduce imperfect information in a centralized school allocation mechanism.

We develop and estimate a model of school application in a large centralized allocation system where students engage into a costly search process to acquire information about schooling opportunities. We show that the tension between the competing explanations in the literature such as the existence of unsophisticated agents or the low valuation of school quality, may be resolved by the existence of imperfect information over schools attributes. Formally, students engage in an iterative and costly search among alternatives to build an endogenous consideration set. A key implication of the model is that schooling decisions are exerted without the full examination of all available options, which may lead to suboptimal decisions.

Our work is related to the recent literature on empirical market designs that is best summarized among others by Fack et al. (2019); Agarwal and Somaini (2018); Calsamglia et al. (2018).[4] While these papers introduce key innovations to the analysis of centralized school allocation systems, our setting differs with the existence of imperfect information, which results from the size of the market. An exception is Kapor et al. (2018), which uses survey data to study the role of beliefs in a centralized mechanism. In contrast to their approach, which focuses on "belief errors", we study the search behavior of students in this setting where information is accessible at cost. This framework allows us to endogeneize how consideration sets are formed. As such, our work is related to a large literature on search frictions that dates back to Stigler (1961).[5] The notion that individuals engage in search over products has been studied in the industrial organization literature as well (Goeree, 2008; Honka et al., 2017; Dinerstein et al., 2018).

The main methodological contribution of our paper is to formulate the school application process as a dynamic search problem. The sequential acquisition of information about schools implies a familiar optimal stopping structure. As such, the decision to stop searching depends on the "luck" of the student, which turns out to be key to rationalize the observation that similar individuals make very different choices and produce credible counterfactual simulations.

We estimate the model using administrative data from Ghana's senior high school

[3]For an exception see Kapor et al. (2018).

[4]The predominant focus in the empirical school choice literature has been on lottery-based admission and the Boston mechanism (see Abdulkadiroglu and Sonmez, 2003). The recent literature tries to quantify the welfare gains associated with changing the allocation mechanism. Related literature also includes Kapor (2016); Abdulkadiroglu et al. (2017); Walters (2018).

[5]A recent literature in decision theory analyzes the role of information attention and rational inattention in individual choice (see Masatlioglu et al., 2012; Caplin and Dean, 2011; Sims, 2003).

choice system in the year 2008, when students were allowed to rank a maximum of 6 programs. 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, measures of mismatch and overtime correlation in school cutoffs.

We show that the key empirical regularities of individual choices in our data (reverting) provides identification content for preferences, beliefs and considerations. Leveraging information on individuals who submit the same set of schools but in different orderings, we isolate the effect of preferences for schools and beliefs about admission chances from that of consideration. Further, we note that as an individual uses the same test score to measure her admission chances in all choices, there exists an inter-dependency between schooling options, which introduces a non-linearity in the value of a ROL.

Our estimated parameters are as expected, indicating individuals' preferences for school quality and its proxy and a dis-utility associated with technical programs. We also find that search costs are high, which implies that consideration sets are relatively small. The median consideration set consists of 10 choices, which implies that several dominated options are considered, and ROLs are not monotonically ranked in admission probability.

Then, we quantify the welfare implications of school choice in the presence of imperfect information on school characteristics. Our analysis of welfare shows that only a quarter of the efficient allocation is realized. As welfare is measured in terms of willingness to travel, our results imply that the cost of boarding could be divided by 4.

Since the large majority of the welfare losses are incurred by low ability students, our findings suggest that the initial objective to increase equity in access to secondary education may be negated. Further computations show that 58.7% of the welfare loss can be attributed to the inability of students to gather information about all alternatives, while the remaining 41.3% is due to uncertainty (test score and coordination frictions).

Finally, since the planner may be more informed about schooling opportunities than students, we study whether restricting choice could be welfare improving. Counterfactual outcomes will depend largely on the information set of the planner. On one extreme, if we were to assume that the planner knows all the parameters of the utility function, the efficient allocation is achieved. A more realistic policy where the planner, knows only that school quality affects utility positively, can increase welfare by 72%.

The rest of the paper is organized as follows. In Section 1, we describe our data and report several empirical regularities. Section 2 describes the model, while estimation and identification are discussed in section 3. The estimation results are presented in Section

4. Section 5 presents a welfare analysis. Finally, section 6 discusses alternative models while section 7 concludes.

1 Motivation

Our data comes from Ghana, where the national school system consists of six years of primary school, three years of junior high school (JHS), and three years of senior high school (SHS). In contrast to most higher income nations, high-school graduation (SHS) is the final degree for almost 80% of students (Duflo et al., 2017). Starting in 2005, students completing junior high school apply for admission to senior high school through a centralized application system. One may wonder about the rationale for organizing a nationwide education system for teenagers. As we will show later, top academic programs are located in few regions, a national education system gives a pathway to elite schools for students in rural locations.[6]

Students apply to specific academic programs within a school and can submit a ranked list of up to six choices. Available programs include agriculture, business, general arts, general science, home economics, visual arts, and several occupational programs offered by technical or vocational institutes.[7]

After submitting their ranked ordered lists, students take a standardized Basic Education Certification Exam (BECE). The application system then allocates students to schools based on a serial dictatorship where priorities are determined by the BECE score.[8]

Students who are unassigned at the end of the algorithm are administratively assigned to a nearby program with remaining vacancies. Our data, which provides individual wishes along with BECE scores as well as admission outcomes, consists of the universe

[6] Student placement before the program was based on yearly regional selection meetings.

[7] For exposition simplicity, we use the term school to refer to a bundle school/program. When strictly referring to a school, we use the term senior high school (SHS).

[8] In practice, the algorithm is implemented as a student-proposing deferred acceptance mechanism, where schools are indifferent between students so the outcome is equivalent to a serial dictatorship. The algorithm is 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.

of junior high schools (grades 6-9) in 2008.[9]

In the remainder of this section, we study in detail individuals' application behavior as well as admission outcomes, revealing some regularities that will guide our modeling strategy.

1.1 Students and Schools.

This section reports the basic statistics behind our data. We describe the characteristics of the students, before considering schools.

The full sample in 2008 consists of 340,823 students, among which, 160,936 students (47%) passed the qualifying exam and are therefore considered for the matching.[10]

Table 1 reports the basic characteristics of students. Over half (53.2 %) are male and the average age is 16.6. Geographically, almost 40% of students are located in the Ashanti and Accra (capital) regions to be compared to 35% for the total population. Student performance on the BECE exam ranges from 185 to 469 points out of a possible 600.[11] As such, students have very heterogeneous chances of gaining admission to any given program. Table 1 also reports that younger male students are over-represented among higher test score students. Similarly, high test score students are over-represented in the Accra and Ashanti regions. In the absence of information on family background, we proxy it using measures of academic success at the junior high school level (average BECE score, and BECE pass rate).

---

[9] We checked the consistency between applications and admission outcomes, and we find a 95% matching rate. The inconsistency comes from students who submit a truncated ROLs. It is possible that the admission office administratively assigned all students with missing choices. We recode the admission outcomes of those students. For approximately 10 students, we had to adjust the size of the program.

[10] Among the 160,936 qualified students, 24 do not apply to any school, 44 apply to only one school, 52 apply to 2 choices, 170 to 3 choices, 8,788 to 4 choices, 8,769 to 5 choices. In total 152,167 (94.55%) of the students apply to all six choices.

[11] There are 281 distinct values of the BECE score. Yet, ties in the matching process are extremely rare because of the size of the choice set. For example the mode of the test score distribution is 262 with 1,324 students. Among the first choice of this group, there 654 distinct options. When ties occur during the estimation, we break them randomly.

Table 1: Students Characteristics

		Students test score (quantiles)			
Characteristics	All	190 – 254	254 – 286	286 – 328	328 – 469
Age	16.648	17.256	16.946	16.524	15.847
Male	0.584	0.573	0.584	0.592	0.585
JHS quality	291.936	259.525	274.648	294.928	339.650
JHS pass rate	0.685	0.496	0.619	0.743	0.886
Regions					
Ashanti	0.233	0.150	0.222	0.281	0.282
Accra	0.255	0.156	0.195	0.255	0.418
Central	0.080	0.110	0.090	0.073	0.047
Eastern	0.099	0.119	0.111	0.095	0.072
Volta	0.063	0.089	0.073	0.058	0.032
Western	0.090	0.113	0.103	0.084	0.061

Notes: The table shows the characteristics of junior high schools students who qualify for senior high school placement. Characteristics are computed for the full sample, and by quantile of student test score. For concision, the largest six of the ten regions are reported.

Then, we consider the other side of the market, which consists of schools. There is a total of 641 schools, and some offer as many as 33 programs.[12] In total, there are 2,113 school-programs.[13] Table 2 reports the characteristics of schools.

[12]This includes traditional high school, and both technical and vocational training institutions.

[13]Our attempts at reducing the dimensionality of the problem have failed as there are no systematic matching patterns between individuals and schools. As such, restrictions on the set of schools or individuals considered may alter the matching outcomes, and limit the scope of any counterfactual analysis.

Table 2: Choice Characteristics

		School cutoffs (quantiles)			
Characteristics	All	158 − 215	215 − 240	240 − 286	286 − 433
Boarding	0.559	0.344	0.411	0.635	0.866
Colonial	0.066	0.005	0.009	0.034	0.259
Religious	0.217	0.135	0.154	0.249	0.278
Size	66.503	75.251	68.200	76.092	82.924
Gender					
Boys Only	0.034	0.002	0.004	0.018	0.136
Girls Only	0.057	0.018	0.014	0.037	0.162
Coed	0.896	0.977	0.979	0.942	0.701
Programs					
Agriculture	0.113	0.097	0.113	0.053	0.011
Business	0.122	0.167	0.165	0.120	0.093
General Arts	0.163	0.170	0.147	0.155	0.174
General Science	0.194	0.195	0.152	0.201	0.243
Home Economics	0.101	0.045	0.093	0.108	0.194
Technical	0.150	0.152	0.172	0.189	0.132
Visual Arts	0.055	0.081	0.082	0.051	0.021

Notes: The table shows the characteristics of all schools/programs. The size of the choice is defined as the number of vacancy reported by the school. The gender category reports the gender exclusivity of the school. Characteristics are computed for the full sample, and by quantile of school selectivity measured by realized cutoffs in 2008. For concision, all the technical programs have been grouped into one category.

There is substantial variation across programs. Over half of the programs (55.9%) offer boarding facilities. The presence of boarding facilities implies that students may gain admission everywhere in the country. The elite schools (6.6%) were established by the British colonial administration before Ghana gained independence in 1957 (colonial), and a little more than 20% of the programs were offered in schools with a religious affiliation. The average program admits 66.5 students, with a range from 10 to 120. While co-education has been generalized over the years, 10% of schools are still single sex, with approximately two thirds of them being girls-only programs. A substantial share of the single-sex schools are also religious, and were established pre-independence.

Finally, General Sciences and General Arts are the most commonly offered programs, accounting for approximately 35% of available options. Technical and vocational education represent 15% of the choices. We now consider the same characteristics by school selectivity. In this setting, selectivity is based on observed cutoffs in 2007.[14] There is a strong correlation between school quality and the indicators for boarding, pre-independence and coed status. That is, a very large majority of high-selectivity schools offer boarding facilities (86.6%), over a quarter of them date back to the colonial era, and single-sex schools are over-represented among them. We also note that although there is not a monotonic relationship between school quality and size, more selective schools appear to offer more seats. Finally, exploring the level of selectivity by programs shows a consistent pattern: choices in general sciences and home economics are the most over represented among high selectivity options. On the contrary, choices in agriculture and visual arts are the least selective.

1.2 Choices.

The matching mechanism is based on serial dictatorship, which is strategy-proof when individuals are allowed to submit an unrestricted number of choices. However, in our case, ROLs are truncated at six (6) choices, which prompts individuals to be strategic. Determining which subset of schools to submit is a complicated problem. Students must find the right balance between sought-after schools, which are likely to be selective while insuring themselves against the risk of unassignment. While there is no simple strategy to construct a portfolio, the literature has provided some results about the properties of the optimal portfolio.

Proposition 1 ***Haeringer and Klijn (2009).*** *Let $N_p = \mathcal{S} = \|U_n > 0\|$ be the set of alternatives with positive utilities. Then, the optimal strategy consists of choosing N among N_p, and ranks them according to the true preference ordering.*

Proposition 1 illustrates a simple property: while finding the optimal portfolio may not be obvious, the ordering within the portfolio is. Specifically, not ranking choices according to true preferences conveys the risk of getting assigned to a less preferred option.[15] In addition to the ordering of choices reflecting true preferences, students are often advised to diversify their rank-order lists, including selective schools as well as safer options. In the rest of this section, we describe individual choices in detail.

Table 3 presents descriptive statistics on students' ranked choices. We report the characteristics of each ranked option to determine whether there exists a consistent pattern across choices. As mentioned before, students were allowed to list six choices in 2008.

[14] We have the same results when using realized cutoffs in 2008.

[15] For example, when choosing between Harvard and Hudson University (fictional and less selective), one should list Hudson as your first choice only if you prefer it. But, if you do, you should not list Harvard, as it is unlikely that you would get rejected at Hudson and admitted at Harvard under a serial dictatorship.

Table 3: Characteristics of the ranked choices

	Choices					
	1	2	3	4	5	6
Colonial						
mean	0.253	0.150	0.102	0.072	0.025	0.017
sd	0.434	0.357	0.303	0.258	0.155	0.130
Religious						
mean	0.248	0.217	0.207	0.200	0.292	0.294
sd	0.432	0.412	0.405	0.400	0.454	0.455
Board						
mean	0.869	0.813	0.759	0.681	0.605	0.582
sd	0.338	0.390	0.427	0.466	0.489	0.493
Coed						
mean	0.744	0.870	0.918	0.947	0.965	0.979
sd	0.437	0.337	0.275	0.223	0.185	0.143
Cutoffs						
mean	318.359	298.838	284.470	269.438	247.247	241.697
sd	59.493	56.065	54.037	52.300	34.286	33.335
Distance						
mean	34.148	32.929	30.680	26.507	30.287	30.904
sd	47.634	45.654	43.708	41.568	28.402	28.351
Programs						
Agriculture	0.057	0.070	0.078	0.092	0.076	0.081
Business	0.196	0.213	0.202	0.187	0.181	0.170
General Arts	0.399	0.389	0.393	0.387	0.392	0.393
General Science	0.138	0.105	0.089	0.079	0.075	0.064
Home Economics	0.098	0.104	0.108	0.113	0.096	0.101
Technical	0.045	0.048	0.051	0.059	0.120	0.127
Visual Arts	0.066	0.072	0.078	0.083	0.059	0.064

Notes: Table shows the characteristics of all schools/programs by ranked choices for 6 choices. Distance is evaluated between the centroid of the junior high school and senior high school districts using GPS coordinates.

Table 3 shows that students are more likely to list a school that was established pre-independence as their first choice. That is, 25.3% of first choices are colonial schools, to be compared to 1.6% for the sixth choice. A similar pattern is observed for schools with boarding facilities, and the coed status. On the contrary, there is a gradient for religious schools only for the first 4 choices: the share of religious schools among fifth and sixth choices is even higher than among first choice. This finding is surprising given the high correlation between school selectivity and the religious status.[16]

Then, we examine the distance between a student's junior high school and selected senior high schools. 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 London to Geneva). At the same time a substantial share of students apply only in their home region. Preferences for distance are convex. Students' first choice are on average 34.1 miles away from their junior high schools and their second choice programs are 1.2 miles closer to them. Their third and fourth ranked choices are 30.7 and 26.5 miles away, but their last two choices are further away at a distance of 30.3 and 30.9 miles on average. Even though there is no clear gradient, the dispersion in distance tend to decrease over choices.

In contrast to preferences for distance, selectivity of ranked programs decreases monotonically. The cutoff score of a students' first choice is 318 but falls to 242 for the lowest ranked choice, which represents a difference of almost 1.5 standard deviations in the BECE score distribution.

Finally, we examine discrete program characteristics and reveal additional characteristics of aggregate choices in table 3. General arts is the most popular program track, with 39 percent of students choosing it as their first and sixth choices, which is mostly explained by the large supply of general arts programs. General science has the steepest gradient in choices. 13.8 percent of students choose a general science program as their first choice to be compared to 6.4 percent as a sixth choice. Preferences for technical programs show the reverse pattern, with 4.5 percent of students choosing one as their first choice to be compared to 12.7 percent for their sixth choice. The remaining programs are more equally represented across choices with 19 percent of students choosing business, 10 percent choosing home economics, 8 percent choosing agriculture, 7 percent choosing visual arts, and 4 percent choosing technical programs.

Selectivity in Choices. After reporting the aggregate characteristics in choices, we provide a deeper analysis of school selectivity in ranked choices. As explained before, students should find a balance between selective schools and include some safety options. In this section, we discretize school selectivity by quantiles, and then report choices by school selectivity in table 4.

[16] There are two types of religious schools in our data. The first consists of colonial era schools, which are mostly single-sex, and while the second is composed of newly established schools, which provides a Koranic or evangelical education. The former are very selective, while the latter are not.

Table 4: School Quality in Ranked Choices

School Selectivity	Choices 1	2	3	4	5	6
(158,215]	0.038	0.058	0.083	0.131	0.151	0.184
(215,240]	0.048	0.072	0.098	0.136	0.179	0.207
(240,286]	0.194	0.261	0.300	0.322	0.457	0.442
(286,433]	0.720	0.609	0.520	0.411	0.214	0.167

Notes: Table shows the distribution of school selectivity among ranked choice. Reading: 3.8% of schools with cutoffs in $(158, 215]$ were ranked as first choice. Observed cutoffs in 2007 are used.

Table 4 shows that approximately 72.0 percent of first choices consist of the most selective schools, a ratio that decreases to 16.7 percent for sixth choices. Conversely, 3.8% of schools ranked as first choices consist of the least selective schools, while this ratio increases to 20.6% for the sixth choice. Yet, there is an odd pattern for schools belonging to the second quartile: respectively 17.9% and 20.7% of school/program reported as fifth and sixth choices are made up of them. One would have expected fewer high selectivity schools in the fifth and sixth choices.

Similarly, the ratio of selective schools among sixth choice may strike as high, but this could be driven by high-ability students. As a consequence, it is possible that *individuals may not be diversifying as they should*.

Diversification. While instructive about aggregate patterns, the former tables do not inform us on the internal consistency of individual choices. We therefore analyze whether individuals target a specific set of characteristics in their application behavior.

Table 5 investigates whether students apply to choices with the same set of characteristics – such as academic tracks, SHS, districts and regions. Our intuition is that individuals may target specific characteristics and in the pursuit of these characteristics, choices may not reflect a thorough trade-off. We find that only 11.8% of individuals apply to a single academic track throughout their entire list, which suggests that the large majority of individuals do not attach a high value to a single academic track. A larger share of students apply to two and three programs (resp. 31.6% and 33.4%). With respect to senior high schools (SHS), individuals almost exclusively apply to multiple SHS, suggesting that there is no attempt to get into a particular SHS, and then switch to a different academic program afterwards. Finally, choices are not scattered geographically, the large majority of students apply to schools in 1 or 2 regions. The latter finding underscores the concentration of top academic tracks in a limited number of regions. This intuition is confirmed by the spread in the number of districts individuals apply to.

Table 5: Repeated Characteristics in Choices

	Number of distinct choices					
	1	2	3	4	5	6
Programs	0.118	0.316	0.334	0.184	0.044	0.004
SHS	0.000	0.004	0.020	0.077	0.231	0.668
Regions	0.557	0.318	0.106	0.017	0.001	0.000
Districts	0.042	0.154	0.297	0.336	0.170	
Colonial	0.112	0.888				
Boarding	0.115	0.885				

Notes: Table analyses whether individuals target specific program characteristics in their applications. Reading: 11.8% of students apply to a single program through out their application.

The observation that students do not attach a strong value to program nor specific SHS points to a genuine attempt to construct portfolios of schools that balance ambition and preferences. Yet, the fact that choices can not be characterized by a limited set of variables suggests that the portfolio construction problem may be complex, with potential substitution between multiple choices. The complexity of constructing a portfolio may lead to mistakes.

Reverting. We report that despite the complexity of the problem at hand, individuals' choices are coherent. On average, cutoffs are decreasing for later ranked choices, as well as other indicators for school quality. Individuals are diversifying their portfolios, including different academic tracks, as well as varied schools in selectivity bandwidth.

In this section, we show that these aggregate statistics conceal insights gained from examining the ordering of choices in individual portfolios. When admission chances are stable across years, ROLs should be such that the most selective is ranked first, then the second selective school, and so on. As such, the absence of this property (reverting) is weakly dominated strategy. This result remains true even though the most selective schools are not the most desirable ones. Intuitively, this implies that some schools may be dominated (lower preference, higher selectivity), dominated options will not be listed if the choice set is large enough. Using our data, we can test for this property.

Table 6: Reverting

		All	Students test score (quantiles) 1^{st}	2^{nd}	3^{rd}	4^{th}
Non-reverters	Realization	0.055	0.016	0.028	0.049	0.133
	Past	0.052	0.015	0.024	0.046	0.127

Notes: Table reports the prevalence of reverting in the panel.

An analysis using the ex-post realization of cutoffs and submitted ROLs suggests that 5.5% of individuals report a rank-order-list where choices are monotonically ordered by cutoffs (non-reverting). This ratio decreases to 5.2% when considering the cutoff of the previous period. The ratio of non-reverts almost doubles when we consider high ability students. The extent of the reverting problem is uncovered using partial ordering. After two choices, 30.2% of the sample have reverted, and increases to 59.8% then 78.1% after the third and fourth choices. In addition, many students revert multiple times. For example, 38.9% of the students who revert at choice 2, end up reverting again between choices 3 and 4. Finally, we find that as many as 45% of students who revert could have been admitted into a more selective school by changing the ordering of choices within their ROL.

Further analysis shows that students who do not revert, have on average a test score of 341 to be compared to 286 for the full sample. Yet the distributions are not disjoint, suggesting that the reverting behavior can not be explained by ability alone. A regression analysis of the determinants of reverting shows the weak correlation to basic observed characteristics (age and gender). However, a junior high school fixed effect is not significant, suggesting that students are not being coached in some junior high schools. Finally, there is a strong correlation between residence in the capital region and reverting. That is, students from the region of Accra make up 38.3% of non-reverters but constitute only 25.1% of the general population.

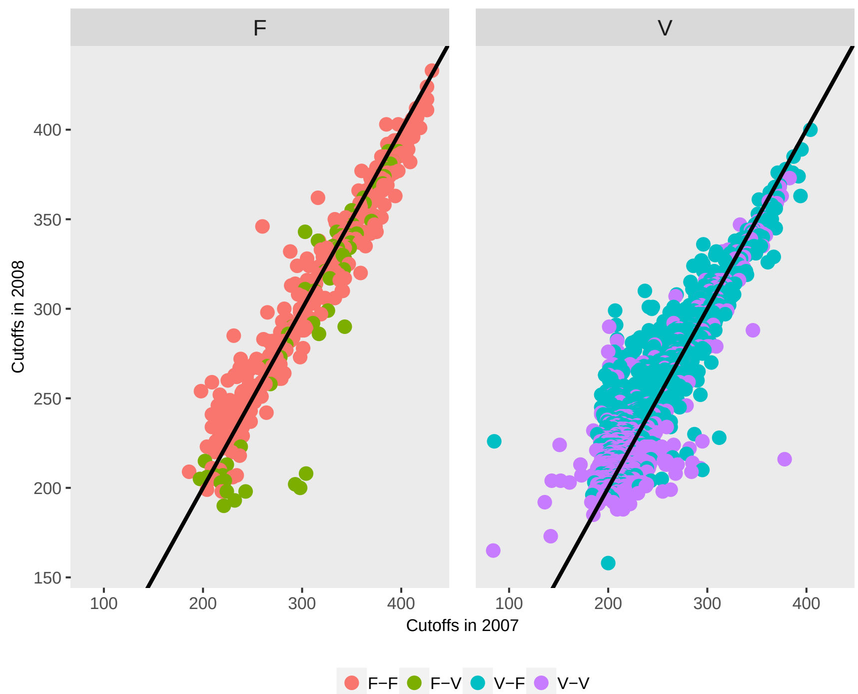
In the next section, we show that there are obviously good reasons for reverting in our setting.

1.3 Uncertainty.

In this section, we introduce the notion of uncertainty. There are two sources of uncertainty in our setting. The first, which we refer to as individual uncertainty, comes from the fact that students 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 about the preferences of other students. The problem of aggregate uncertainty is exacerbated by the fact that prior to 2008, there have been several changes in the institutional background. In addition, in contrast to other countries (Tunisia for example), it appears that no information about cutoffs and schools is provided to participants. Since our analysis is based on administrative data, we do not have any information about beliefs. However, we can check whether the selectivity of schools is constant across years.

Figure 1 reports a strong correlation between cutoffs across years (0.84). Yet, the correlation drops to 0.37, when we account for schools with at least one opening. Nonetheless, we note that there is more variation in cutoffs for lower selectivity schools.

Figure 1: Stability of Cutoffs



Notes: Cutoffs are defined as the test score of the last individual admitted, regardless of capacity. "F" stands for full, while "V" stands for Vacant. The top panels are conditioned by the vacancy status in 2007. The heatmap is weighted by the number of vacancies, which gives the impression that there more vacant schools.

Figure 1 also documents variation along the dimension of not reaching capacity. That is, we show that while the observed cutoff in 2007 is a good predictor of vacancy status, there is substantial variation among the schools that were full in 2007 and vacant in 2008, and vice versa. As such, uncertainty is not only related to cutoffs, but also to the existence of a non-negligible share of schools who will remain vacant.

It is worth nothing that transitions from Full to Vacant and from Vacant to Full are accompanied by significant changes in cutoffs. That is, the large majority of schools that make a transition from full to vacancy (71%) are located below the 45 degree line, which implies that their cutoff (as measured by the test score of the last admitted student) decreases. Similarly, we observe a reverse pattern for schools that transition from vacant to full. Also, the large majority of choices who remain vacant throughout the two years have lower cutoffs. As a consequence, we view cutoffs as measured by the score of the last admitted student to contain information about the relative selectivity of the school.

In conclusion, uncertainty is a key characteristic of the market we analyze. As the ranking of schools over years is not constant, reverting may be optimal.

1.4 Mismatch.

Finally, we consider the outcome of students' application behavior. Table 7 reports the placement outcome of students, and shows that the large majority of students gain admission into their first three choices. That is, 27% of individuals are admitted into their first choice. Interestingly, not only high test score students are placed into their first choice, more than 15% of the low test score students are assigned to their first choice as well, which speaks to potential non diversification in the ROLs of some students.

Table 7: Placement

		Students test score (quantiles)			
Placement	All	1^{st}	2^{nd}	3^{rd}	4^{th}
1	0.272	0.150	0.218	0.273	0.453
2	0.198	0.129	0.169	0.223	0.274
3	0.184	0.159	0.192	0.216	0.168
4	0.160	0.188	0.189	0.184	0.077
5	0.023	0.055	0.024	0.010	0.003
6	0.018	0.043	0.019	0.007	0.001
Administrative	0.145	0.275	0.189	0.086	0.024

Notes: Table reports the placement of students, including administrative assignment. The placement outcome is also reported by student test score quantiles.

The value of the fifth and sixth choices appears relatively limited, as only around 4%

of students get admitted into those choices. This observation is at odds with the share of students who were unassigned (administrative assignment), which is 14.5.[17] Administratively assigned students are placed into a local school. [18] As expected, administrative assignment is closely related to test score - inasmuch as 27.5% of lowest ability students are administratively assigned, while only 2.4% of the highest ability students end up unmatched. There are very few matching setting in the world with two digit mismatch rates. The combination of administrative assignment and the observation that fifth and sixth choices are not well utilized suggests that the aggregate trends in the data may conceal some application shortcomings. In addition, the high share of first choice admission suggests that it may be possible to improve the allocation of some of the matched students.

Table 8 shows that only 55.78% of the schools end-up at capacity.[19] Not surprising, the vacancy rate is decreasing in school selectivity. However, vacancies are not confined to low selectivity schools. That is, only 76.4% of the 25% most selective schools are at capacity, a ratio that increases to 79.5% when we consider the 5% most selective schools. While the median high selectivity school has one remaining seat, least selective schools have more vacancies.

[17]A recent press article states that "About Sixty-seven thousand three hundred and eighty two (67,382) students who qualified for senior high schools (SHS) could not be placed under CSSPS this year (2018)", which suggests that the problem is still present, and may have worsen as there is less than twenty-five thousands (25,000) unmatched agents in 2008.

[18]As we have reported that students are willing to travel to attend good schools, one may view this policy as being counter-intuitive. However, matching in distant schools raises concerns about the ability of parents to pay for boarding.

[19]The imbalance between the number of vacancies and the number of students is driving this number up.

Table 8: Vacancies

(a): Prevalence of Vacancies

			School cutoffs (quantiles)			
		All	1^{st}	2^{nd}	3^{rd}	4^{th}
Cutoffs	Share	0.442	0.669	0.638	0.174	0.236
	Seats (median).	29.000	30.000	32.000	33.000	1.000

(b): Characteristics of vacant schools

		Non vacant	Vacant
Cutoffs	mean.	267.658	230.114
	sd.	55.265	46.598
Colonial	mean.	0.104	0.039
Size	mean.	78.027	72.593
Boarding	mean	0.709	0.374

Notes: Table reports the occurrence of vacancy at the school level in panel a. Panel illustrates the characteristics of vacant and non vacant schools.

In addition, we show that schools with vacancies are larger and less likely to have boarding facilities. The existence of vacancies, administrative assignment along with the high share of reverting suggest a deeper problem, which we posit to be the existence of imperfect information. We view imperfect information as a better explanation because of the existence of administrative assignment, vacancies and reverting. For example, the existence of unsophisticated agents alone can not rationalize the double digit share of administratively assigned students since unsophisticated agents tend to apply to schools within reach.

Under the hypothesis of imperfect information, students are not aware of the characteristics of the schools but are required to engage into a costly search to acquire information about school characteristics. The existence of vacancies implies that the conditional probability of being accepted in a more selective school after being rejected from a less selected school may not be zero. In essence, this implies that dominated options may be listed as part of the dynamic search problem for schools.

2 A Model of School Application Under Imperfect Information.

In this section, we develop an empirical model, which is consistent with the key facts presented in section 1. To that end, we introduce frictional search in the standard school application problem.[20] We formulate the school application process as a search problem, where students iteratively acquired information about schools. The assumption that search is sequential implies that the decision to stop acquiring information depends on the "luck" of the student, an assumption that is likely to help us match the observation that some high ability students do not list 6 choices. In addition, the search framework allows us to generate mismatch on the extensive (administrative assignment for students and vacancy at the school level) and intensive (matching into more desirable school) margins. Finally, the existence of search cost may compel students to consider only a subset of choices, leading to sub-optimal decisions.

Framework. The school choice problem is summarized as follows. A finite set of students $\mathcal{I} = \{1, 2, \ldots, I\}$ apply to a finite set of schools $\mathcal{S} = \{1, 2, \ldots, S\}$.

Each school has positive capacity, and students can opt out of the matching system and enjoy an outside utility u_0, which is set to 0 for simplicity. A student is characterized by a set of observed attributes $\mathbf{x}$ and a test score t which is unknown when she submits a rank order list (ROL). The latter defines individual admission priorities while the former captures preferences. Schools have an observable set of characteristics given by $\mathbf{z}$, and a fixed capacity denoted by K. Finally and following the literature, we assume that each school has a cutoff q.[21]

The assignment mechanism is a serial dictatorship, with student priorities determined by test scores. Students submit a rank-order list that does not have to reflect their true preferences over schools. In our current setting, students can submit up to six choices, a constraint that makes it even more likely that rank-order lists may not reflect true preferences. The payoffs depend not only on which schools are listed, but also on the order they are listed in. We assume that students understand this.

We assume that students act as price takers, taking admission probability as given.[22]

[20] A recent literature provides a theoretical foundation to search as originating from endogenous consideration sets under the notion of rational inattention. Our approach is related to theoretical models that study the implications of rational inattention for choices using search technology (Masatlioglu and Nakajima, 2013; Caplin and Dean, 2011).

[21] Formally, letting q_j be a cutoff : the minimal test-score required for admission at school j is defined such that $\mathbb{1}\{\mathcal{D}_j(q_j)\} \leq C_j \quad \forall j \in \mathcal{J}$ where $\mathcal{D}_j()$ is the demand for that school j with capacity C_j. Our application does not enforce this definition, as we assume that there is a positive probability that an student gets admitted to a school where a past cutoff is higher than his realized test score.

[22] see for example Azevedo and Leshno (2016); Agarwal and Somaini (2018) among others for a similar assumption.

While restrictive, this assumption is required for tractability.

Preferences. The utility for an individual with characteristics $\mathbf{x}$ matched with a school with attributes $\mathbf{z}$ is given by $u(\mathbf{z}, \mathbf{x})$. We follow Berry and Pakes (2007), and assume that the indirect utility function includes a disturbance term ϵ that is additive and separable between school attributes and student characteristics:

$$u(\mathbf{z}, \mathbf{x}) = \gamma \mathbf{z} + \Gamma \mathbf{z}\mathbf{x} - d(\mathbf{z}, \mathbf{x}) + \epsilon \tag{1}$$

where the set of school attributes, $\mathbf{z}$, includes school quality as measured by the average score of students admitted the previous year. Ideally, one would use a measure of value-added for school quality, however, we do not have any such variable. Additional controls include school size, and indicators for boarding facilities, pre-independence, religiosity and program track. The set of individual characteristics, $\mathbf{x}$, consists of realized individual test score, gender, age, and proxies of family background measured at the junior high school level. $d()$ provides the distance between the student and school locations. Since over 99 percent of programs are public schools, we use distance as our numéraire. As a consequence, the parameters γ measure student's willingness-to-travel for each school attributes, while Γ capture the interactions between students and school characteristics. Finally, ϵ is idiosyncratic tastes for schools. Students know their tastes, which are unobserved by the econometrician. The error term ϵ is i.i.d and follows a distribution $\mathcal{N}(0, \sigma_\epsilon)$.

Beliefs about admission chances. At the time of submitting lists, priorities and cutoffs are not known. Priorities are based on individual test score obtained from a national exam that will take place 4 months later. In the next section, we describe in more detail how students learn about cutoffs. For now, we state that students do not know cutoffs, but may learn cutoffs (q_j) through search.

We assume that before the exam, each student has a prior about his test score, which is private information τ, not observed by the econometrician. We assume that agent i forms beliefs about realized test score following $t_i = \tau_i + \varepsilon_i$ with the cdf of ε given $F_\varepsilon(\cdot)$, and the variance of ε is a population-wide parameter. Essentially, we assume an error term ε that captures deviation between subjective and objective probabilities. Given q_j, a student with t_i can construct her admission probability as $Pr(t_i > q_j)$.

Finally, we introduce an additional effect of uncertainty. As reported by figure 1, cutoffs may be not stable over years, which implies that the ranking of schools may not be either. The introduction of an error term in the formation of beliefs about admission chances captures this effect. However, uncertainty will also affect the conditional probability of being rejected from a seemingly less selective school and being admitted to a seemingly more selective school. When considering two schools with cutoffs q_l and q_n, this conditional probability is denoted by $\sigma(q_l, q_n)$.

Search. In order to acquire information about school characteristics, students engage in a sequential and costly search among the alternatives.[23] One could imagine a process where a student visits a school and gathers information.[24] As a consequence, we view frictions as emerging from the existence of a large number of options.

Since our analysis of applications, presented in section 1, did not reveal any systematic pattern on search directness, we assume that search is random. While this assumption may be strong for high ability students, the technology of directed search may be extremely hard to identify given our data.

In order to describe the search problem, we resort to the notion of a consideration set, which allows us to dissociate the search process from the construction of a ranked-order-list. Through search, agents build and expand a consideration set, denoted by $\mathbf{c} \subset \mathcal{S}$. The existence of imperfect information gives rise to the notion of a consideration set, whereby only a subset of available alternatives will be considered for choice.

Each draw corresponds to a school. One element of the consideration set is a couple $(\mathbf{z}, q)$. At the beginning of time, the consideration set is empty $\mathbf{c} = \emptyset$.[25]

Individuals search through available options, at cost $c(n)$, where $n = \|\mathbf{c}\|$, is the size of the consideration set. Furthermore, we assume that $c(n)$ is positive and given by $c(n) = c \times n$. In our empirical application, we include a stochastic shock ξ in the cost of application, which helps generating heterogeneity in the size of ROLs. The stochastic shock is iid across individuals and follows a distribution $F_{\xi}(\cdot)$.

The value of search to a student with test score τ, characteristics x, and a size n consideration set $\mathbf{c}$ is given by $\mathcal{V}_n(\tau, \mathbf{c})$:

$$\mathcal{V}_n(\tau, \mathbf{x}, \mathbf{c}) = \sum_{s \in S \setminus \mathbf{c}} \mathcal{U}(\tau, \mathcal{B}(\mathbf{c}_s))p(s) \tag{2}$$

[23] There is a literature that studies the nature of search in environment characterized by costly acquisition of information. While most applications are in consumer search, the theoretical foundation explores questions related to sequential versus non sequential search, and the impact of recall on consumer choices Morgan and Manning (1985); Morgan (1983). The main difference with our current setting is that individuals search in order to construct an optimal portfolio.

[24] This is a purely theoretical concept. If that was the case, and search cost was increasing with distance, we would expect students to list fewer distant schools, which is not the case.

[25] The assumption of empty initial consideration is arguably strong. One could imagine that students may acquire information about a number of schools from parents, friends and school teachers. However, an analysis of the applications does not reveal that students from the same junior high-schools apply to the same set of schools, and we do not have any information about the information set of parents. As a consequence, we opt for this strategy. In addition, this strategy is observationally equivalent to set an initial set of random schools.

where $\mathbf{c}_s = \mathbf{c} \cup (\mathbf{z}_s, q_s)$. $\mathcal{U}(\tau, \mathcal{B}(\mathbf{c}))$ is the highest utility attainable for a student with test score τ and consideration set $\mathbf{c}$. Finally, $p(s)$ is the probability of drawing a school s, which under random search is uniform. Note that searching for another option comes with cost $c(n+1)$.

Optimal behavior. Given this structure, we can characterize the optimal strategy of the student. First, we describe some properties of the value function.

Lemma 1 *The value from search $\mathcal{V}_n(\tau, \mathbf{x}, \mathbf{c})$ is increasing and concave in n with initial value equation given by the expected value $V_1(\tau, \mathbf{x},) = \sum_{s \in S} Pr(\tau_i \geq q_s) u(\mathbf{x}, \mathbf{z}_s)$.*

Lemma 1 implies that the value of search is increasing. In expectation, larger consideration sets yield higher utilities. Concavity induces that the marginal gain of search is decreasing, which compels individual to stop searching.

Proposition 2 *Assume $V_1 > c_1$, then there exist a unique reservation value $V_n^\star$.*

Proposition 2 implies that there exists a unique reservation value, a point at which search is not profitable. Prior to any search, the structure of the value function implies a distribution of consideration sets based on observed heterogeneity. The realized distribution may differ based on the specific draws.

Portfolio construction. Finally, we consider how individuals can construct the best set of schools $\mathcal{B}(\mathbf{c}')$ given $\mathbf{c}'$. Since individuals use the same test score to evaluate her admission chance throughout the search process, choices are interdependent. Explained differently, rejection in the first choice conveys additional information on one's test score and the expected distribution of cutoffs. Recently, Shorrer (2019); Calsamglia et al. (2018) have proposed a method to construct the best set of schools in that setting. In our case, it turns out that consideration sets are relatively small, and as a consequence, we can compute all the combinations and pick the one that yields the highest utility as in He (2012).

3 Estimation

Our final sample consists of the 169,097 individuals who complete the BECE exam. Although the individuals who do not pass the exam could provide additional information about test score uncertainty, we opt against this strategy since their test scores are reported as missing. In order to limit the number of available choices, we do not consider schools that were not subscribed at all, which leaves us with 2,113 choices. We estimate the model by Simulated Method of Moments. 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:

$$\mathcal{L}(\Theta) = -\frac{1}{2}\big(\widehat{m} - m(\Theta)\big)^T \widehat{W}^{-1}\big(\widehat{m} - m(\Theta)\big) \tag{3}$$

where $\widehat{m}$ is a set of empirical moments, and $\widehat{W}$ is the weighting matrix.[26]

In the rest of this section, we provide identification argument for the parameters for preferences, beliefs about admission chances and search cost. Then, we describe our moments.

3.1 Identification

The basic identification problem is related to the fact that a school may not be listed because it was not considered or was considered but not selected.[27] As such, choice data alone can not separate these two channels. In addition, disentangling preferences from private information (beliefs about admission chances) can be difficult in matching mechanisms that use exam scores to allocate students.

The standard identification argument in discrete choice with limited attention involves using data on consideration sets (Jolivet and Turon, 2014; Honka et al., 2017; Dinerstein et al., 2018) or restrictions that some determinants of attention are orthogonal to preferences (Barseghyan et al., 2019; Heiss et al., 2016).

In the rest of this section, we argue that data on rank-order-list provides identification content for all the parameters of the model. The intuition is to leverage variation in the ordering of choices across individuals. Specifically, as ROLs provide multiple listings for each individuals, across individuals comparisons allow us to isolate the role of preferences, beliefs and consideration. We argue that identifying utility shocks, preferences for schools, and beliefs about admission chance does not require knowing the consideration set of students. That intuition is developed in the rest of this section.

Proposition 3 *The variance of the preference shock σ_ϵ is identified.*

[26]We use a diagonal weighting matrix, with the elements set equal to the inverse of the diagonal variance-covariance matrix of the empirical moments. Since we have discrete dependents, approximation of the gradient vector are sensitive to the chosen step size. We therefore calculate the derivative by first approximating the function by a low-order polynomial function as we vary each parameter locally.

[27]See Flinn and Heckman (1982) for an analysis of the identification problem, which consists of separating large utility shock from search frictions. The key result of Flinn and Heckman (1982) is that a "recoverability" condition is required, which implies that the untruncated distribution can be recovered from the truncated distribution.

Proof Consider a subset of individuals with same observed characteristics $\mathbf{x}$ who apply to the same set of schools $\mathcal{S}_1$ but provide different orderings. Let $u(\mathbf{x}, \mathbf{z}) = \bar{u} + u_\epsilon$. As u_ϵ is independent of $\mathbf{x}$ and $\mathbf{z}$, utilities shocks are iid across individuals, then var(u_ϵ) identifies σ_ϵ.

Proposition 3 exploits a feature of our data, and the intuition is as follows. Consider two individuals with same observed characteristics $\mathbf{x}$ and test score t. Further assume that those students report the same set of choices, but with a different ordering. Since the individuals report the same choices, we can ignore any variation that may be due to differences in consideration sets. Also, as choices are the same across individuals, we can further discard any variation due to preferences and beliefs. As a consequence, variation in the ordering of choices across similar individuals identifies the utility shock. We should note that it may be difficult to observe several individuals who report exactly the same set of 6 schools with different ordering. However the proof does not require that the complete ROLs to be similar across individuals. The same identification argument applies for any subset of two choices within the ROL.

Proposition 4 $\mathbb{E}u(\mathbf{x}, \mathbf{z})$ *is identified.*

Proof Consider a subset of individuals with same test score t who apply to the same set of schools $\mathcal{S}_1$ but provide different orderings. As u_ϵ is known, $\mathbb{E}(p(t)u(\mathbf{x}, \mathbf{z}) \mid \mathbf{x})$ is identified.

The identification of expected utility builds on the previous result. To see this, consider two individuals with same observed test score t, but with different observed characteristics ($\mathbf{x}$). Further assume that those students report the same set of choices, but with a different ordering. As before, since the individuals have similar choices, we can ignore variation due to differences in consideration sets. Since the variance of the preference shocks has been uncovered, expected utility is identified. As we exploit variation in individual characteristics, identification requires to observe similar choices along the full support of choice characteristics $\mathbf{z}$.

As the test score is included both in preferences and admission chances, separating $p(t)$ from $u(\mathbf{x}, \mathbf{z})$ requires additional sources of variation. Separating preferences from beliefs is a daunting problem. Manski (2004) advocates to collect additional data on beliefs as a choice data may be compatible with multiple beliefs. In the absence of such data, we rely on the structure of the choice problem to provide identification content.

Proposition 5 *Preferences* $u(\mathbf{x}, \mathbf{z})$ *and admission chances* $\mathbf{p}(t)$ *are separately identified.*

In order to establish that preferences and beliefs are separately identified, we exploit a feature of choice problem. As a student uses the same test score to evaluate her admission probabilities across all choices, admission chances are inter-dependent across alternatives. Specifically, the value of a ROL $(1,2)$ is given by $p_1U_1 + (p_2 - p_1)U_2$, where p_j and U_j are admission chances and utilities of choice j. As a consequence, preferences and beliefs enter non-linearly in the value of the choice. While non-linearity is usually viewed as an artificial source of identification, we argue that it is a generic property of any choice problem when the decision maker chooses multiple products with correlated winning probabilities. As such, it is not only an essential feature of the problem at hand, but also the main source of difficulty in constructing the optimal portfolio.

Next, we can consider the identification of consideration set. Given preferences and beliefs about admission chances, it is trivial to show that variation in the choices of students with the same observed characteristics $\mathbf{x}$ and test score t, that cannot be justified by the existence of an utility shock, will identify the consideration set. Essentially, assume that we observe two individuals with the same observed characteristics making different choices. From propositions 5 and 3, the deterministic and stochastic components of utility are known. As such, differences that can not rationalized by utilities will identify the consideration set.[28]

Finally, we should note that $\sigma(q_m, q_n)$, which is the conditional probability of being accepted into a more selective school after being rejected from a less selective school is not identified. Intuitively, this non-identification may be explained by the fact that the mechanism of reverting does not operate through the conditional probability but hinges on the existence of dominated options in the consideration set. As a consequence, we calibrate this probability to 0.002, which corresponds to the value that minimizes the moment criterion. This calibration does not have any impact on preference parameters but does appear to have a very modest effect on the scale of the uncertainty parameter.

The next section builds on this identification result to construct moments.

3.2 Moments

We construct empirical analogs that capture the identification content provided by the data. Mainly, we use two sets of moments. That is, for any simulated portfolio, $\mathcal{S} = \{\mathcal{S}_n\}_1^6$, we evaluate:

[28] We should note that 9,068 of our students report less than 6 choices (truncated list). For those students, both the consideration set and choices are observed. While we do not use this information directly for identification, we can assess whether our model is able to replicate the consideration sets of those students.

1. Expectation of schools' observable characteristics by ranked choice

$$\mathbb{E}\left(Z_{ij}(\mathcal{S}_{i,n})\right)) \tag{4}$$

2. Conditional expectation between students' and schools' observable characteristics by ranked choice

$$\mathbb{E}\left(Z_{ij}(\mathcal{S}_{i,n})|X_i\right) \tag{5}$$

We compute these moments for the full sample by ranked order lists. Then, consistent with our identification strategy, we compute these moments for any pair of choices, which are ranked by multiple individuals. We differentiate the cases where the choices are ranked in the same ordering or in different ordering. To be sure, this is extremely expensive combinatorial analysis that requires changing six choices into multiple unordered pairs. These unordered pairs are then clustered by the two distinct orderings (decreasing and increasing in orderings). Then, our main moments ($\mathbb{E}(\mathbf{z})$ and $\mathbb{E}(\mathbf{z} \mid \mathbf{x})$) are evaluated on these groups.[29]

We use additional moments. Consistent with the search literature, we match the vacancy rate, and the share of administrative assignment. Then, we match the ratio of monotonically ranked portfolios to ensure that the model is able to generate reverting as well as the decreasing average school quality over ranked choices. Then, we match the share of individuals reporting less than 6 choices in their ROLs. This moment provides a good benchmark for observed consideration sets. Finally, we match the correlation between past and realized cutoffs, which helps capture the effect of nonlinear. In practice, this moment pins down the distribution of ε that reconciles subjective and objective admission chances.

Given the set of school characteristics and the number of moments, the model is overidentified. As a consequence, we target a limited number of characteristics, and gauge the out-of-the sample validity of our model using moments that are not included in the estimation. Notably, we omit the distance variable, and many programs characteristics. Further, we match only the choices for the first, third and fifth choices, such that fitted moments on the second, fourth and sixth choices provide additional evidence of out-of-the sample validity.

3.3 Estimation

Since some of the moments depend on the matching outcomes, we solve the search problem for all individuals, then construct the matching allocation. The stochastic components

[29] Section B in the appendix enumerates the number of unordered pairs.

are integrated-out through simulations. In our final specification, we have 137 moments and 28 parameters. We estimate the model using POUNDERS (TAO implementation), which is a derivative-free model-based algorithm for nonlinear least squares.

4 Results

This section presents our estimation results. First, we discuss preference estimates, then search cost parameters and the implied consideration sets. Beliefs about admission chances are then presented. Finally, the fit of the model is discussed.

Preferences Table 9 presents the parameters governing the utility of students.[30]

Our final specification includes several school characteristics, which are then interacted with key individual characteristics.[31] Estimates for school characteristics are consistent with qualitative evidence, presented in section 1. On average, students prefer boarding schools, and older schools established before Ghana gained independence. Religious and single-sex schools appear to impact negatively students utility although the effect is not significant. Students have a significant preference for general sciences and general arts programs and a strong negative taste for technical programs.

Compared to these average preferences, students with higher test scores place more emphasis on programs in general sciences and in pre-independence schools. Male students have a stronger preference for school quality and value less older schools relative to females, with a significantly weaker preference for boarding schools.

Students from higher-performing junior high schools place relatively more value on school quality and value less boarding facilities, but have a weaker preference for older schools.

[30] We omit the intercept in presenting these parameters.

[31] Boarding is an indicator variable for the existence of boarding facilities at the school level, while Colonial, Religious and Coed, are indicator variables for the creation of the school prior to Ghana independence, religious school, and mixed gender education. G. Science, G. Arts, and Technical are program indicators for General Sciences, General Arts and Technical. Score is the standardized test score obtained on the BECE, which is re-scaled between 0-1 for the estimation. Quality is measured as the average test score of students admitted in the program the previous year. Finally, J.Quality and J.Rate are demographic characteristics at the junior high school level capturing respectively the average score and the passing rate at the BECE.

Table 9: Estimation results (utilities)

Variables	Est.	Std.
Boarding	2.737	0.215
Colonial	2.065	0.415
Religious	-0.211	0.264
Coed	0.833	0.231
G. Science	1.881	0.119
G. Arts	0.163	0.118
Technical	- 1.142	0.311
Quality	0.502	0.091
Score × Colonial	1.093	0.321
Score × Boarding	0.379	0.183
Score × Quality	0.399	0.132
Male × Colonial	0.644	0.281
Male × Boarding	-0.222	0.371
Male × Quality	0.589	0.329
J. Quality × Colonial	0.148	0.141
J. Quality × Boarding	0.251	0.46
J. Quality × Quality	0.729	0.198
J. Rate × Colonial	1.093	0.117
J. Rate × Boarding	0.326	0.231
J. Rate × Quality	0.895	0.482
Score × G.Science	0.205	0.091
Score × G.Arts	-1.319	0.377
Score × Technical	-5.638	0.356

Notes: The table shows parameter estimates under our preferred specification. A description of the variables is provided in the footnote 31.

Search and consideration sets Then, we quantify the role of search cost. The costly nature of search implies that all choices are not considered as such this parameter along with the search technology is the main driver of consideration set size. As we parametrize $c(n) = c \times n$, the parameter c can be interpreted as the cost of considering an additional school. Our estimate is approximately 0.073 to be compared to the average (resp. median) utility of 3.2 (resp. 3.88). The shock to the search cost (ξ) has mean 0 and a standard deviation of 0.012, which implies that 16.4% of the search cost is stochastic.

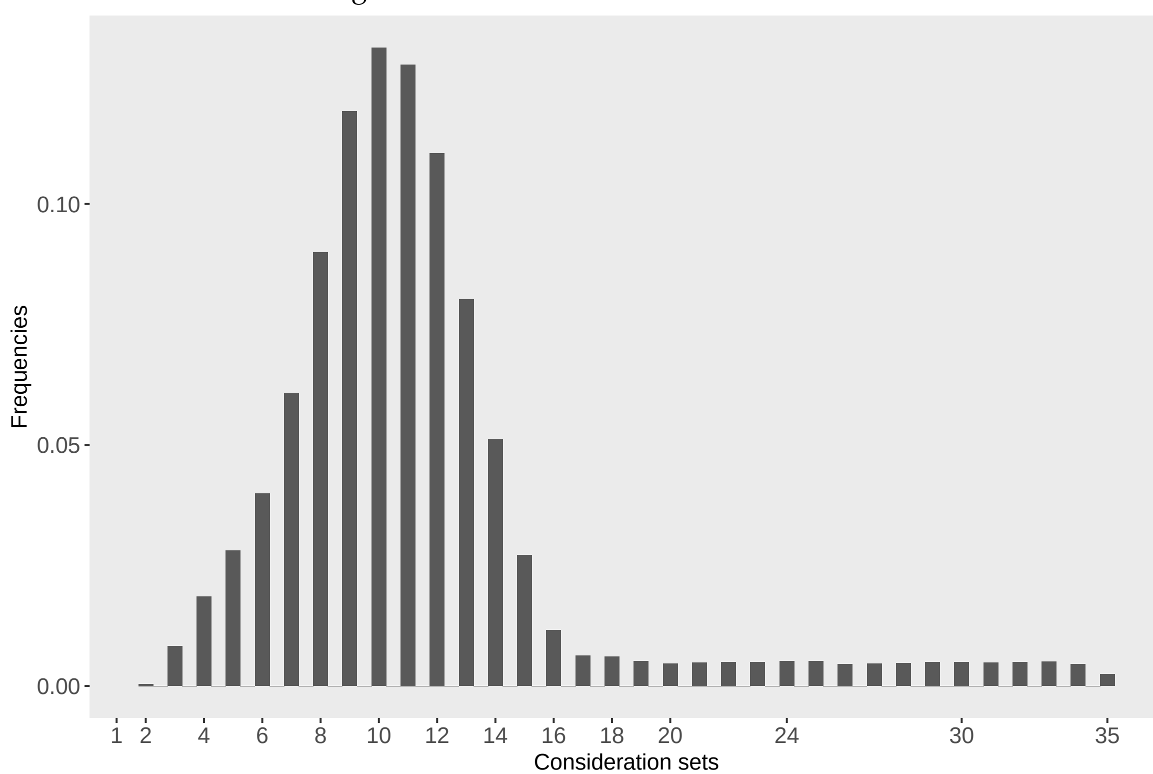
Table 10: Estimation results (cost and shocks)

Variables	Est.	Std.
c	0.073	0.026
σ_{ε}	0.207	0.045
σ_{ξ}	0.012	0.002
σ_{ϵ}	0.195	0.027

Notes: The table shows estimates of the marginal cost of search, and the shock parameters.

The main implication of these costs is related to the size of consideration sets. Our findings suggest that students consider between 2 to 35 choices. Figure 2 represents the implied distribution of consideration sets under our model. Interestingly, we match almost perfectly the share of individuals who are reporting truncated lists of 3, 4 and 5 choices.

Figure 2: Size of the consideration sets

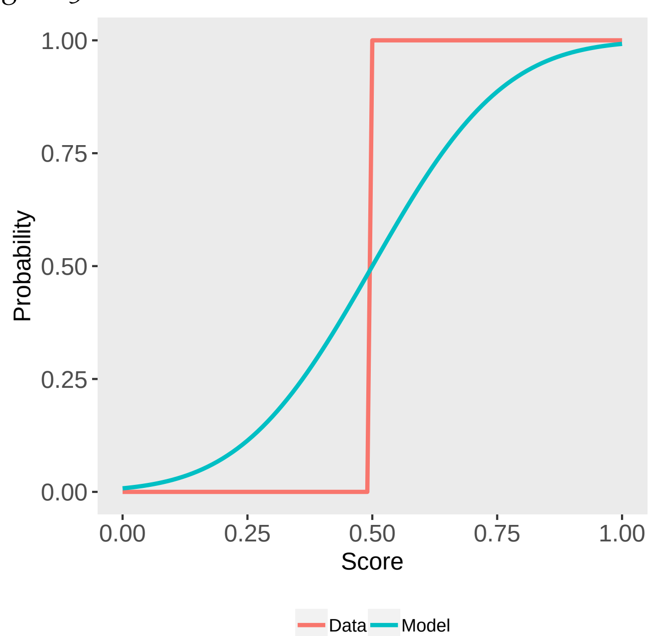


Notes: Distribution of the size of consideration set in the estimation.

Figure 2 shows that the very large majority of students (84.8%) consider between 6 to 14 schools. On the left side of the distribution of consideration sets, we denote 5.3% of students whose lists are truncated, while the right tail (9.9%) consists almost exclusively of students belonging to the top decile of the test score distribution (consideration set between 15 to 35 choices). The consideration set of the median student contains 11 schools, which is negligible when compared to the choice set of 1,182 choices. The inability to direct search pushes the value of search down when the number of choices is large. As we will document in section 6, this feature of the model is essential to rationalize the matching patterns observed in the data. To summarize, these findings are consistent with our data as students are likely to submit dominated options when their consideration set is small.

Beliefs and admission chances Finally, we consider the parameters that characterize students' beliefs about admission chances. Since we do not have any data on beliefs, we introduce an error term that ensures that realized admission probabilities coincide with prior beliefs. As such, σ_ε captures the level of uncertainty in the matching process, whose effect is reported in Figure 3.

Figure 3: Beliefs and Realized Admission Probabilities



Notes: Admission probability for a school with a cutoff of 0.5 for a grid of test score ranging from 0 to 1.

Figure 3 represents the admission chance to a school with a cutoff of 0.5 for a grid of test score ranging from 0 to 1. Under the normal error assumption, admission chances are smooth on the support. As such, the admission probability of lower ability students is over-estimated, while that of high ability students is under-estimated.

Goodness of Fit Since we are interested in counterfactual simulations, we present evidence on how well our model fits the data. We simulate ROLs under the model and compare them to the real data.

Figure 4: Fit

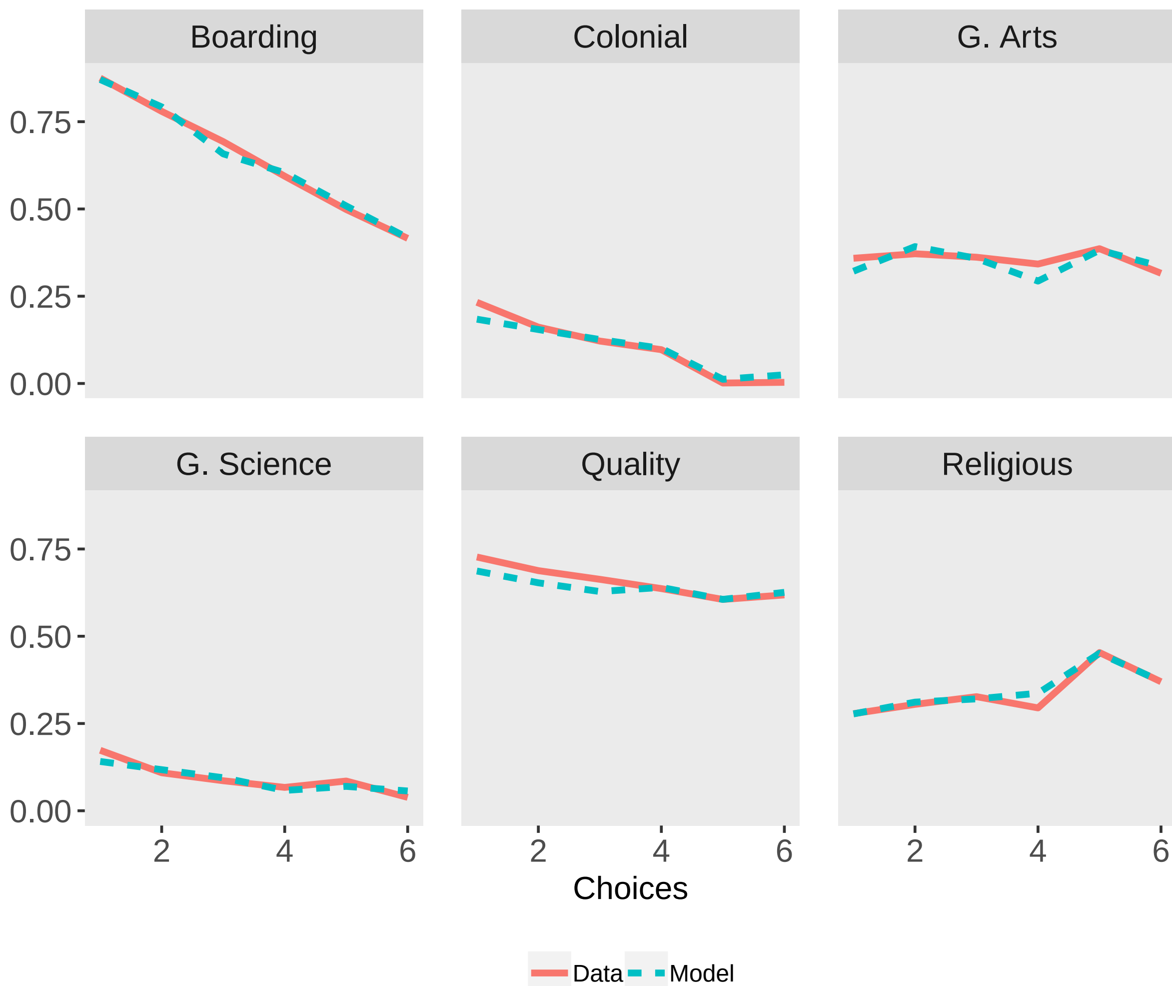
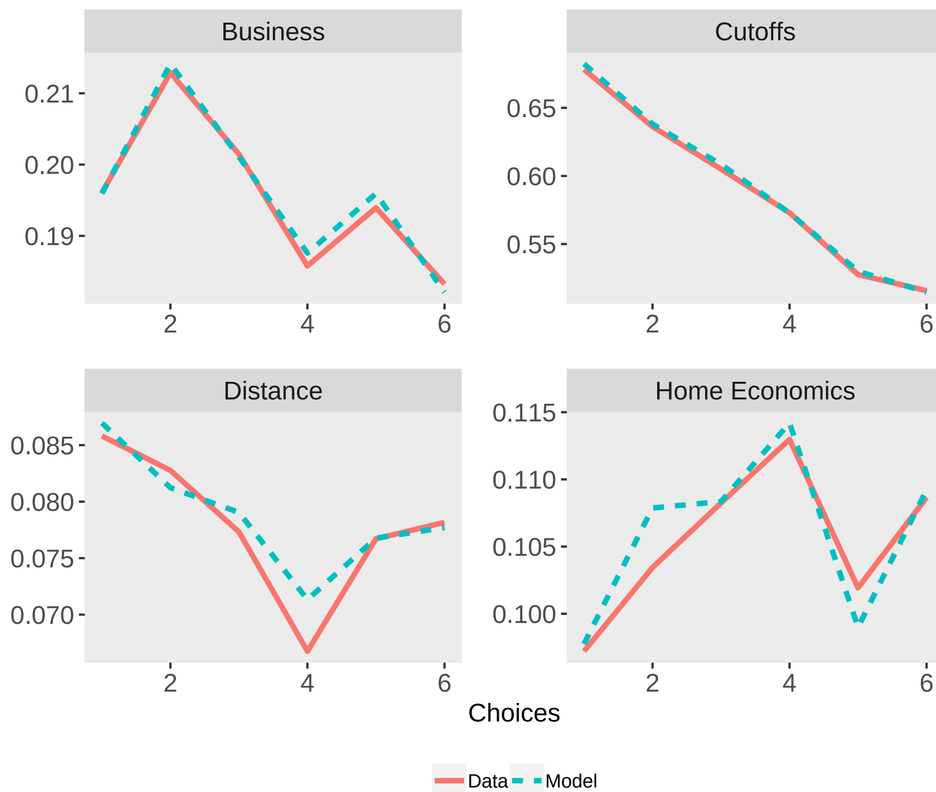


Figure 4 presents the fit for a subset of moments targeted in the estimation. For concision, we focus on a limited number of moments although in the main text, while the complete fit is presented in appendix A. We should note that the estimation targets only choices 1, 3, and 5. Our model fits the distribution of choice characteristics for each decision very well. The model captures the sharp decline in school quality over ranked choices. The actual and predicted school profiles (namely availability of boarding facilities, the creation of the school prior to independence, and religiosity) are very close. The patterns of academic program choices are accurately predicted as well. General arts is more popular than general science in both cases. In all cases, we match not only the patterns across choices, but also the actual value of the variables. Figure 5 presents additional evidence on out-of-sample fit, using data on moments not used in the estimation. Inasmuch as these characteristics are not completely correlated with other characteristics used in the estimation, the non targeted moments provide out-of-sample validation.

Figure 5: Non targeted Moments



Notes: Out of the Sample Validation. Cutoffs has been divided by 3 to produce variables that are on the same scale.

The patterns observed in the targeted moments largely hold with the non-targeted moments. However, when variations across ranked choices are not monotonic, the model produces changes that appear to be sharper across choices. Yet, even for the distance variable which does not appear to be fitting very well, there is never a difference of more than 0.01 between true and estimated moments. To sum up, for a model, which is over-identified, the out-of-sample fit is very good, suggesting that the model fits the patterns observed in the data both quantitatively and qualitatively. As such, we are confident that we can use the framework to perform counterfactual simulations.

5 Counterfactual Simulations

In this section, we analyze the efficiency content of our model. Given individual preferences, and technological constraints on vacancies, we investigate whether the planner can achieve a better allocation. Then, we build on this result to explore ways to improve the allocation of students.

5.1 Quantifying inefficiencies

Inefficiencies stem from the existence of search friction and uncertainty.[32] Efficiency, or its lack of is clearly reflected in the existence of vacancies, administrative assignment, and potential mismatches among the matched students. Our findings in section 1 suggest that a substantial share of students could have been admitted to a more selective school/program by changing the ordering.

We quantify the respective importance of search frictions, and uncertainty on individual welfare. We propose a simple utilitarian welfare function $\sum_i U_i^*$ that aggregates individual utilities, where U_i^* is the utility individual i derives from the school he was assigned. Individuals who do not gain admission into any of their choices are randomly assigned to a local school.

We consider three settings. Frictional Application is our benchmark case. Then, the problem of the Constrained Planner (CP) is analyzed. The CP maximizes total welfare subject to preference and technology constraints. Since matching is centralized, we assume that the planner, upon observing realized test score, can alleviate frictions.

Finally, we evaluate welfare assuming there are no frictions, such that individuals can construct Optimal Portfolios (OP). In OP, individuals observe perfectly the characteristics of all choices, and hence can optimally select the ROL that yields the highest utility. Recently Shorrer (2019); Calsamglia et al. (2018) have proposed a method to recover such a portfolio. The idea is to use dynamic programming to account for the inter-dependence in

[32] Another source of inefficiency comes from the discrepancy between private and social values of search which arises because of the standard "overcrowding" among students: when an extra person lists a school, it reduces the availability of vacancies for other students. An externality that students are not likely to internalize. See Abdulkadiroglu et al. (2015) for a theoretical analysis of this problem in school choice. However, we do not address this. Also, the fact that students can submit only 6 choices, induces additional efficiency concerns.

admission chances across choices. We implement this strategy. In doing so, one potential problem is related to the fact that cutoffs may change endogenously as students use a different method to construct portfolios. As a consequence, we solve for the Bayesian Nash Equilibrium. That is, given an initial set of cutoffs denoted by q^0, and preferences estimated in our setting summarized by u, we solve for the following algorithm.

1 Individuals select the rank-order lists that maximize expected utility.

2 Given submitted lists, students get admitted to schools, and the realized matching determines the new distribution of cutoffs.

3 Repeat until cutoffs converge.

Table 11 reports the results.

Table 11: Efficiency

	Benchmark	Constrained Planner (CP)	Optimal Portfolios (OP)
Utilities	100	397.5	233.4
Cost	100	-	31.3
Admin Assignment	0.16	-	0.09
Vacancies	0.52	0.11	0.47

Notes: Under CP, the planner assigns students based on her knowledge of individual preferences using realized test scores as priorities. Under OP, individuals submit optimal ROLs, which are used as input in the matching. Benchmark is normalized to 100.

In interpreting these results, it is not obvious how to treat the cost component of search. That is, under the constrained planner, there is no search per se. As a consequence, we focus on utilities, and our results may be viewed as a lower bound on the welfare gain. We show that the constrained planner achieves approximately four times more welfare than our benchmark. As we measure welfare in terms of willingness-to-travel, our results indicate that students could travel four times less for school, and the cost of boarding could be divided by four in the economy.

The gap between the frictional application and the allocation achieved under the constrained planner highlights the importance of inefficiencies. Interestingly, we find that eliminating search frictions would multiply welfare by 2.3.[33] As a consequence, we can conclude that 58.7% (233.4/397.5) of the welfare loss can be imputed to the existence of frictions, while the remaining 41.3% is due to uncertainty (test score and coordination frictions).

[33] Which still implies that students are still able to construct optimal portfolios.

5.2 Choice Paradigm and Welfare

School choice is based on the premise that students (or parents) know better which school to attend. The standard paradigm in choice theory, *the more options the better*, reinforces the notion that expanding the horizon of choices, beyond an assigned neighborhood for example, improves welfare. This is the rationale that motivates a national placement system in Ghana.

However, when school decisions are made without the full examination of all available options, students may be worse off. Furthermore, as the number of choices increases, it becomes almost impossible for decision makers to know all choices. One extreme case of restricting choice is the efficient allocation, where the planner has full information on the preferences of all students has information on the preferences of all students. In this section, we analyze whether, restricting choice in a realistic fashion could be welfare improving.

In the first experiment, we let the planner assign students under the assumption that individuals value only school quality. As a result, the planner will assign the highest test score student to the most selective school. The second experiment considers also preferences for programs. The intuition for doing so is that many high achieving girls apply to home economics programs. Using programs in the assignment allows us to assign less high achieving boys to those programs, without taking any stance on the intrinsic "value of home economics" to boys.

We let the planner assign a student to a less selective program, if the program is more popular and the difference in cutoffs between the two choices is less than 20 points.[34] We measure the popularity of academic tracks using an over-supply index, which is defined as the difference between the supply of academic tracks and the share of academic tracks among first choices. Results are described in Table 12.

[34] This represents a very rough approximation of the importance of programs in individual utility.

Table 12: Restricting Choices and Welfare

	Benchmark	Quality	Quality + Program
Utilities	100	172.2	213.4
Vacancies	0.52	0.29	0.27

Notes: Under quality, the planner reduces individual utility to school quality alone, and assigns the best student to the most selective school. Under quality + program, the planner still assigns the best student to the most selective option, but is allowed to make a trade-off between school quality and program popularity.

We find that a planner, who assigns the highest test score student to the most selective school, increases welfare by 72%. The fact that we produce such a level of welfare gain after reducing the utility function to a single component is yet another sign for the level of inefficiency in our setting. In addition, restricting choices is likely to help lower ability students since they are the group that are the most affected by inefficiencies. The welfare is still substantially lower than the efficient allocation. We also find that welfare more than doubles when the planner allows for substitution between school quality and programs.

6 Alternative Models

As noted before, we made several simplifying assumptions to ensure tractability. This section documents how relaxing key modeling assumptions, namely endogenous consideration set and random search, affects our ability to reproduce descriptive characteristics of the data and our policy conclusions.

6.1 Exogenous consideration set.

Arguably, one of the key implications of the model is that consideration sets are relatively small. Small consideration sets lead to sub-optimal decisions, and as such may be driving our results about inefficiencies. In this section, we consider a model, where the size of consideration sets are exogenously set. We opt for a strategy where consideration set size is homogenous across individuals.[35] Using such a model, we hope to illustrate which features of the data drive the relatively small consideration set, and how conclusions might change when the size of the consideration set increases. We estimate four

[35] Alternatively, we could devise a two-stage model, where students sequentially draw a size then a consideration set.

additional models where the size of the exogenous consideration set takes value 30, 50, and 100. Finally, we consider the full attention model. Results are reported in Table 13.

Using these models, we report three variables, which capture welfare (sum of utilities), and administrative assignment and reverting rates.

Table 13: Welfare Under Exogenous Consideration Sets

	Endogenous	Consideration sets			
	Consideration	30	50	100	All
Utilities	100	105.7	108.2	113.9	114.2
Reverting	94.4	81.2	65.9	53.5	51.2
Admin Assignment	14.8	14.3	13.7	13.1	12.9

Notes: Exogenous consideration sets are set to 30, 50 and 100. "All" corresponds to the full consideration model. Welfare is normalized to produce comparable figures across the different models. Admin assignment refers to the share of students who are administratively assigned.

Welfare increases with the size of consideration sets. For example, when the exogenous consideration set consists of 30 options, welfare increases by 5.7% relative to our benchmark model. However, the welfare gains are not large: the full consideration models leads to 14.2% increase relative to the endogenous consideration model. The relative stability of welfare can be attributed to two mechanisms.

First, as consideration sets become larger, the model needs additional forces to cope with the observed level of reverting. The parameters adjust such that the preference component for school quality and its proxies appear weak, which leads to less dispersion in utilities. Yet, that force alone is not sufficient to generate enough reverting, as illustrated by the steady decline in the share of reverting.

Second, although consideration set size increases, students can submit only 6 choices. As a consequence, uncertainty regarding admission outcomes generates substantial administrative assignment, which is illustrated by a 12.9% rate of administrative assignment under the full attention model.

Finally, the outcome of our policy experiment on restricting choice does not change substantially. When the planner assigns the best student to the most selective school/program, welfare increases by 72% to be compared to 69% and 65% respectively when we impose a consideration set of thirty (30) and fifty (50) schools.

6.2 Directed Search.

Our second extension considers a model, where students can direct their search toward specific schools. A full directed search model, would be difficult to identify given our data and the descriptive evidence presented on the contents of ROLs. In order to circumvent this, we follow Avery et al. (2014) and group schooling options into 3 categories: reach, match and safety. Each of those types is characterized by a distribution, which is known to the students. Then, we iteratively let students choose an optimal type that depends on its characteristics and its current consideration set, upon which she receives a draw of a specific school/program.

The classification of choices into reach, match and safety options turns out to have important consequences on perceived uncertainty. We use a bandwidth parameter to guide this classification. For example, a bandwidth parameter of 20, implies that a student with test score t will be "match" for a school with cutoff $[t-20, t+20]$, a school with cutoffs $(0, t-20]$ will be "safe", while schools with cutoffs $(t+20, +\infty]$ will be "reach". We estimate the model with three different bandwidths (10, 20 and 30). Results are reported in Table 14.

Table 14: Welfare under Directed Search

	Endogenous	Bandwidth		
	Consideration	10	20	30
Utilities	100	95.7	103.8	99.1
Adm Assign	14.8	17.6	13.2	14.6

Notes:

Welfare estimates are in line with our benchmark model. We find that under such a model of directed search, welfare gains are tied to administrative assignment. The main problem of the setting is related to the classification of choices, which conveys a sense of safety that gets tested by the application behaviour of agents. That is, "safe" schools get oversubscribed, which changes their initial status. This phenomenon leads to unstable matching outcomes. A more convincing way to estimate this model, will be to make cutoffs endogenous, which turns out a very difficult undertaking.[36]

Under this model, assigning the best student to the most selective school leads to 71% welfare gain for a bandwidth of ten (10).

[36] In addition to the computational challenge, an iterative algorithm that feeds initial cutoffs, and then updates them after each set of application, is not guaranteed to converge.

7 Discussion and conclusions

This paper develops and estimates a model to understand individual preferences for schools in a large matching market. We introduce imperfect information in the standard school application problem. Search allows students to learn about the characteristics of schools and build endogenous consideration sets. As such, our model allows us not only to match the choices of high ability students, but also to understand how mismatch between students and schools emerges in a centralized allocation system.

We exploit variation across individuals who submit the same of schools but in different ordering to isolate the effect of consideration sets from that of preferences. Our application in Ghana shows that indeed search costs are high, which implies a high level of inefficiency. An analysis of welfare shows that approximately a quarter of the efficient allocation is realized under the current mechanism. Further computation shows that 58.7% of the welfare loss can be attributed to the inability of students to gather information about all alternatives, while the remaining 41.3% is due to uncertainty (test score and coordination frictions). Finally, we show that a policy that would restrict choice in that setting would be welfare increasing – a planner who is only interested in assigning the best student to the most selective school increases welfare by 72%. Alternative models show that our conclusions are robust to relaxing the assumption of random search and exogenously setting larger consideration sets.

Our findings raise new questions about school choice in large matching markets. Given the number of choices, students may not be aware of all schooling opportunities. The size of the choice set imposes several challenges on low ability students. The first is related to the lack of social ties that allow students to easily collect information about schools. Second, liquidity constraints are likely to affect those students more severely. Finally, the expected gain from search is relatively small for lower ability students, which decreases their value to search.[37] Restricting choice in large markets can help mitigate some of the inefficiencies, but clearly not all of them as it recovers only 18% (0.72/4) of the loss welfare. Future works could speak to the optimal size of a matching market.

The methods used in this paper provide several avenues for future research. Although our analysis focuses on key features of the education system in Ghana, the potential behavioral implications can be extended to many countries, as well large school districts in the US. We show that ranking within ROLs is enough to identify consideration sets from preferences. One key extension would be to augment these types of administrative datasets with surveys on beliefs and search behavior to get a better understanding of the nature of search. These extensions are left for future work.

[37] Unfortunately, our data does not allow us to quantify the respective importance of these channels.

References

Abdulkadiroglu, A., Agarwal, N., and Pathak, P. A. (2017). The welfare effects of coordinated assignment: Evidence from the new york city high school match. *American Economic Review*, 107(12):3635–89.

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., Che, Y.-K., and Yasuda, Y. (2015). Expanding "Choice" in School Choice. *American Economic Journal: Microeconomics*, 7(1):1–42.

Abdulkadiroglu, A., Pathak, P. A., and Roth, A. E. (2005). The new york city high school match. *American Economic Review*, 95(2):364–367.

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

Agarwal, N. and Somaini, P. (2018). Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444.

Agarwal, N. and Somaini, P. (2019). Empirical analysis of school assignment models.

Avery, C., Howell, J. S., and Page, L. (2014). A review of the role of counseling, coaching, and mentoring in students? postsecondary outcomes. *College Board Research Brief*.

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.

Barseghyan, L., Coughlin, M., Molinari, F., and Teitelbaum, J. C. (2019). Heterogeneous Choice Sets and Preferences. Papers 1907.02337, arXiv.org.

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

Calsamglia, C., Fu, C., and Güell, M. (2018). Structural Estimation of a Model of School Choices: the Boston Mechanism vs. Its Alternatives. NBER Working Papers 24588, National Bureau of Economic Research, Inc.

Caplin, A. and Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6(1).

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–59.

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

Fack, G., Grenet, J., and He, Y. (2019). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review*, 109(4):1486–1529.

Flinn, C. and Heckman, J. (1982). New methods for analyzing structural models of labor force dynamics. *Journal of Econometrics*, 18(1):115–168.

Goeree, M. S. (2008). Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074.

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

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

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

Heiss, F., McFadden, D., Winter, J., Wuppermann, A., and Zhou, B. (2016). Inattention and Switching Costs as Sources of Inertia in Medicare Part D. NBER Working Papers 22765, National Bureau of Economic Research, Inc.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: evidence from the u.s. banking industry. *The RAND Journal of Economics*, 48(3):611–646.

Hoxby, C. M. and Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review*, 105(5):514–17.

Jolivet, G. and Turon, H. (2014). Consumer Search Costs and Preferences on the Internet. IZA Discussion Papers 8643, Institute of Labor Economics (IZA).

Kapor, A. (2016). Distributional effects of race blind affirmative action. In *Working Paper, Yale University*.

Kapor, A., Neilson, C. A., and Zimmerman, S. D. (2018). Heterogeneous beliefs and school choice mechanisms. Working Paper 25096, National Bureau of Economic Research.

Lucas, A. M. and Mbiti, I. M. (2012). The determinants and consequences of school choice errors in kenya. *The American Economic Review*, 102(3):283–288.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

Masatlioglu, Y. and Nakajima, D. (2013). Choice by iterative search. *Theoretical Economics*, 8(3).

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5):2183–2205.

Morgan, P. and Manning, R. (1985). Optimal search. *Econometrica*, 53(4):923–944.

Morgan, P. B. (1983). Search and optimal sample sizes. *The Review of Economic Studies*, 50(4):659–675.

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

Shorrer, R. I. (2019). Simultaneous search: Beyond independent successes.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy*, 126(6):2179–2223.

A Fit

This section presents the fit for the complete set of moments used for the estimation. Figures 6 and 7 report conditional moments while table 15 present the fit for the moments based on the outcome of the matching. Overall we find that all moments fit very well the data.

Table 15: Fit (moments based on matching)

Variables	Data.	Model.
Adm. Assignment (%)	14.5	14.8
Vacancy rate (%)	44.2	44.3
Share of non-reverting (%)	5.5	5.7
Share of truncated lists (%)	5.1	5.2
Correlation past and realized cutoffs	0.402	0.398

Figure 6: Fit (1)

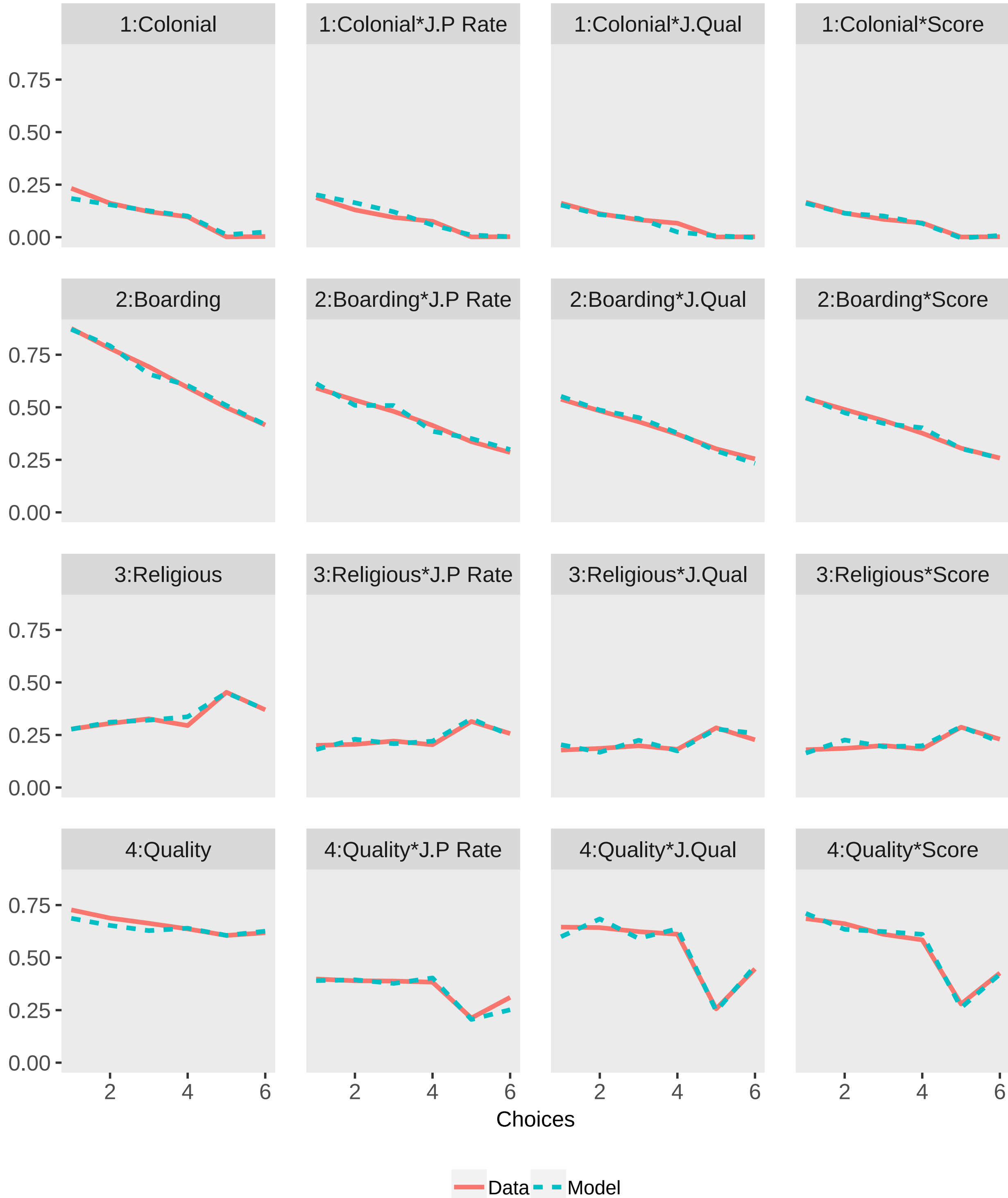
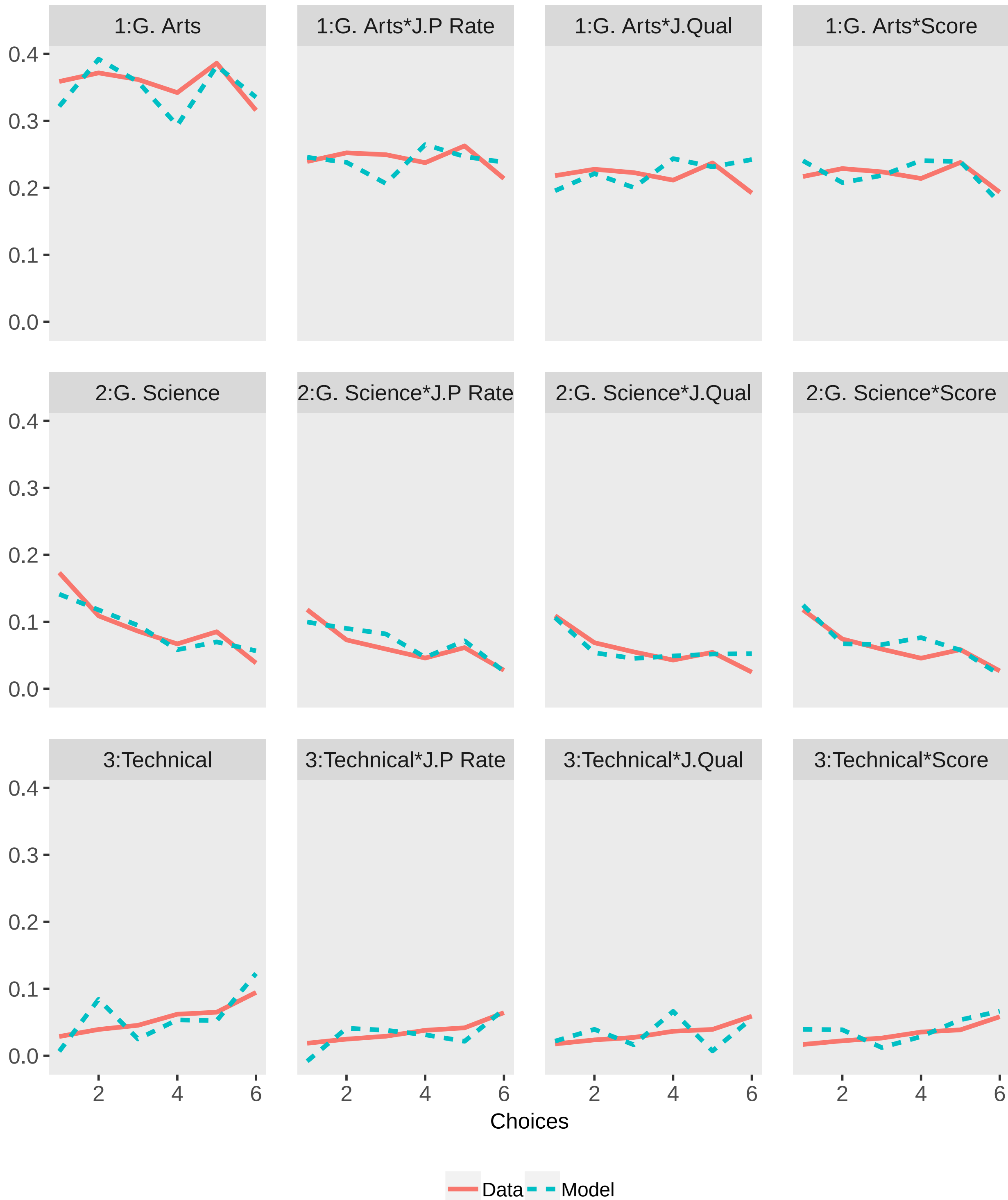


Figure 7: Fit (2)



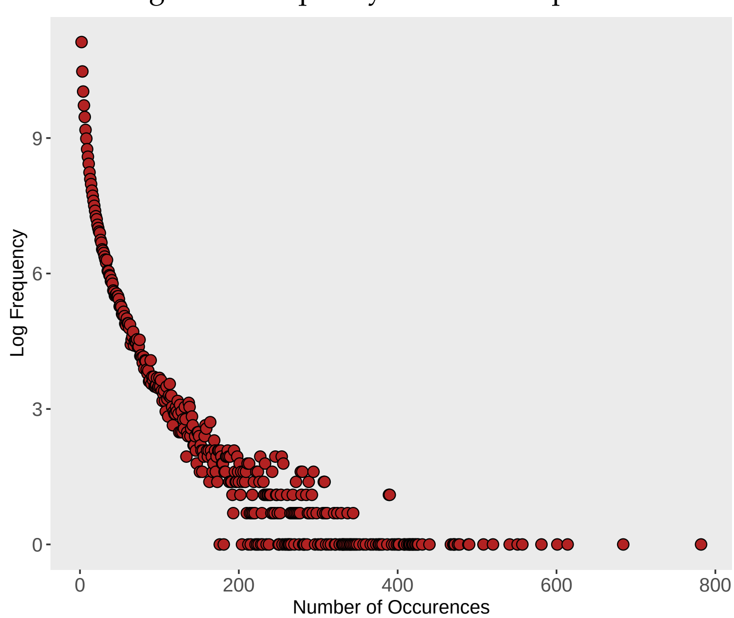
B Identification

As explained before, our identification strategy exploits the fact individuals might submit the same set of schools, but in different order. Once conditioned on the appropriate set of controls, the difference in ordering provides identification content for the utility shocks and preferences. In this appendix, we document whether individuals do indeed report same choices with different ordering.

Our final sample consists of 2,113 choices, and we consider all possible choices. Out of 4,462,656 potential unordered pairs, 430,696 couples are listed at least once (9.6%). Among the 430,696 listed unordered pairs, 197,206 couples are listed once, which leaves 233,490 ordered pairs that can be used for identification.

Figure 8 reports that there is 67,931 unordered pairs that are listed twice. Similarly, there is more than 35,400 unordered pairs that are listed three times. While the number decreases rapidly, we have more than 1,607 unordered pair that are listed more than 100 times. Finally, as suggested by the long right tail of the distribution, some unordered pairs are very popular.

Figure 8: Frequency of observed pairs



Notes: Reading: there are 67,931 = log(11.12) unordered pairs that are listed twice. There is 1 pair that is listed 800 times.

Then, we analyze whether these unordered pairs are provided in the same ordering. We find that 27.2% of those list are provided in different ordering, which leaves us with 63,665 pairs to be used for identification. 2,090 schools are listed in those pairs.