The value of information in centralized school choice systems

Margaux Luflade∗†
University of Pennsylvania

Abstract

Strategy-proofness and assurance of a fair matching are desirable qualities for school choice mechanisms. Although these qualities are theoretical properties of the unrestricted-list deferred acceptance mechanism (DA), it has proven to be hard to go without list-size restrictions in practice. This paper shows how a simple modification to the restricted-list DA, in which students are provided more information about vacancies and offered higher-value options in the event of rejection, can mitigate uncertainty and yield matches very close to what would be obtained under the unrestricted-list DA. I estimate an application-portfolio choice model using administrative data from Tunisia, where a sequential implementation of the DA is used to assign high-school graduates to universities. This sequential implementation creates quasi-experimental variation that allows to separate the identification of students’ preferences for programs from their expectations about their admission probabilities. Counterfactual simulations show that the average student’s expected ex post utility is significantly lower when assignments are made using the standard restricted-list DA than what it would be in the unrestricted-list DA match. Using a sequential implementation, as is done in Tunisia, instead of the standard restricted-list DA, can reduce this welfare loss by 90%.

∗Email address: mluflade@sas.upenn.edu
†I am extremely grateful to Meryam Zaïem whose insights about the context, help with obtaining the data, and encouragement have been critical to the realization of this project. I am grateful to Arnaud Maurel, Peter Arcidiacono, Joe Hotz, and Matt Masten for help and suggestions, as well as continuous support and encouragement. I thank Atila Abdulkadiroğlu, Eli Liebman, Pedro Mira, Derek Neal, Aviv Nevo, Seth Sanders, Modibo Sidibé, Petra Todd, and the participants at the Duke Labor and Microeconometrics workshops and at many other seminars, for very helpful comments and discussions. I thank the Ministry of Higher Education of Tunisia for providing the data. All errors are mine.
1 Introduction

Assignment mechanisms are widely used around the world to decide on students’ admission to public schools and universities. Some have advocated for mechanisms, like those based on the Deferred Acceptance algorithm (henceforth, DA; Gale & Shapley, 1962), that give students incentives to truthfully apply to their most preferred academic programs as a way to pursue “transparency, fairness, and equal access to public facilities” (Abdulkadiroğlu, Pathak, Roth, & Sönmez, 2006). As long as applicants can apply to all programs in their choice set (Haeringer & Klijn, 2009), the DA maintains the incentives for truthful reporting of preferences (Dubins & Freedman, 1981; Roth, 1982) and guarantees that the final match between schools and students is fair in the sense that any school a given student prefers to her match was filled by applicants with higher priority than her (Abdulkadiroğlu & Sönmez, 2003; Balinski & Sönmez, 1999). Virtually all centralized school systems, however, restrict the number of programs students can apply to, and there is little empirical evidence on the consequences of these restrictions on application decisions and on the final student-school matches.

In this paper, I empirically examine students’ application portfolio choice problem when students are not able to apply to all academic programs in their choice set. When they can only apply to a subset of programs, students face the possibility of being rejected from all their listed choices. To avoid rejection, students need to choose their application portfolio strategically, taking into account not only their preferences, but also their probability of being admitted. Students’ expectations about their admission probabilities and the value of their fallback option are then a crucial determinant of where they apply and are ultimately accepted. Taking restrictions on the number of applications as a constraint, I show how an easy-to-implement modification to the DA, in which students are provided more information about vacancies and offered higher-value options in case of rejection, can offset the adverse consequences of list-size restrictions.

I use administrative data from Tunisia, where college applications and assignments are made at a nationwide level using a very simple variant of the restricted-list DA. The Tunisian assignment mechanism provides both a basis for my identification strategy, and a guide for the policy experiment. In the Tunisian system, a score based on performance at a national end-of-high-school exam is used to determine priority at all post-secondary programs, and students know at the time of applying the score of the marginally-admitted student to each program in the previous year. The Tunisian procedure is sequential. The cohort is divided into three tiers based on priority score. Students in the top tier first submit application lists and are assigned via the DA. The vacancies remaining after this first phase are publicly revealed before second-tier students apply. The application/assignment/update process is then repeated, sequentially, for the second and third tiers. In addition, students rejected from all their choices are offered the option to reapply with the next tier of applicants.

This three-phase implementation yields a quasi-experimental setting that I exploit in several ways. First, I provide regression-discontinuity evidence that students change their application choices when provided new information about school vacancies. Second, I recover students’ preferences for programs and test whether students behave strategically when applying under a restricted-list DA, rather than simply listing their most-preferred alternatives. Then, I estimate an application portfolio-choice model and use a counterfactual analysis to evaluate the effects of enabling students to update their expectations about their admission chances and about the value of their fallback option. In line with recent survey findings (e.g., Kapor, Neilson, & Zimmerman, 2018), the model allows students to form expectations about their admission chances that may differ from true probabilities.

The empirical analysis faces two main challenges. The first is an identification challenge generally faced by the empirical literature on matching mechanisms. The mapping from students’ preferences to their application choices depends on their expectations about their admission chances to the different schools.

---

1There is evidence that policy makers are reluctant to cut out this constraint, although they value strategy-proofness and stability of the match. See for instance Pathak and Sönmez (2013) and this recent quote from Roth (2015): “In my description [...] students can list as many schools as they like. We economists recommended that students be allowed to do just that, but on this important detail we did not prevail. So New York City students today can list only up to twelve programs among the hundreds that the city offers. Students who want to list more than that face a strategic choice of which twelve to list.”
With most school applications datasets, the econometrician cannot separately identify students’ preferences and expectations about admission probabilities (Agarwal & Somaini, 2018). I show that the quasi-experiment induced by the Tunisian procedure helps to circumvent this problem because the sequential design induces a subset of students to truthfully report their most-preferred programs. Given the information revelation about vacancies before each group submits applications, the highest-priority student in each group is faced with making a choice under perfect information. It is a dominant strategy for her to list in her application ranking her most-preferred programs among those declared not to be full. Because students may apply to up to ten programs and because most programs have more than one vacancy, not only the first-ranked student, but a subset of applicants at the top of each group have incentives to reveal their preferences. Preferences for this subset of applicants can then be recovered within a standard discrete choice framework; and I argue that, given the setup of the Tunisian mechanism, the utility parameters recovered from this subset of truthful students actually represent the preferences of the full sample. In a second step, I characterize students’ expectations about their admission chances as those rationalizing other students’ observed application lists.

The second challenge to the empirical analysis is computational. Given a student’s preferences and expectations about her admission chances, solving the application portfolio choice model—that is, finding the expected-utility maximizing ordered list of up to ten programs among more than 600 alternatives—is demanding. The computational burden of standard simulation-based methods to estimate models with unobserved heterogeneity is prohibitive. I estimate the model of expectations formation by combining a two-step method-of-moments approach, in the spirit of the estimators proposed by Bajari, Fox, and Ryan (2011) and used by Nevo, Turner, and Williams (2016), with numerical integration techniques, as those illustrated in the context of maximum likelihood by Heiss and Winschel (2008). As in Bajari et al. (2011) and Nevo et al. (2016), and following Ackerberg (2009), the two steps allow me to treat separately the issues of solving the model and optimizing the objective function. I implement the two-step method by using the assumed distribution of unobserved heterogeneity (here, the distribution of expectations-formation types) to write the empirical moments as a mixture of conditional moments. An important distinction from these two papers is that my setting requires allowing the distribution of unobserved heterogeneity to depend on a set of (continuous) observed individual characteristics—e.g. priority and distance to updated information. I use insights from numerical integration to handle this dependence and devise an estimation method in which the model only needs to be solved a finite number of times, outside the optimization routine.

In the counterfactual analysis, I evaluate the effects on students’ assignments of implementing the restricted-list DA sequentially, with various numbers of phases. As a benchmark, I consider the perfect-information setting in which students know exactly what programs would admit them, and that yields the same match as the fair and strategy-proof unrestricted-list DA. The perfect-information case is a limit $N$-phase “fully sequential” scenario, where $N$ is the total number of students in the cohort, and where information about vacancies is publicly updated after every single assignment. Bearing in mind that the time required to implement this “fully sequential” benchmark would be prohibitive in most settings, I then show how a sequential implementation with as few as three to five phases, while being easy to implement, can generate outcomes very close to what would obtain under the perfect-information setting.

I compare students’ expected ex post utility, which they derive from their assignment, under different counterfactual scenarios. Under the single-phase implementation, expected welfare is lower than under perfect information by the equivalent of increasing the average student’s distance to her benchmark perfect-information assignment by 70.3km (43.6 miles). As a reference, the median distance traveled in the data is 100km and the average utility decrease between a student’s most-preferred program and her second (resp. third) most-preferred program is 58km (resp. 93km). The three-phase sequential implementation used in Tunisia reduces by 90% the loss in welfare induced by the use of a single-phase restricted-list DA in an environment where students face uncertainty about their admission chances, relative to a perfect-information setting.

Interestingly, students at all priority levels experience ex post-utility gains from switching from a one-phase to a multiple-phase implementation, but distinct mechanisms drive the gains to high- vs low-priority students. Low-priority students’ gains result from an information-revelation channel. Under the
standard one-phase restricted-list implementation, 6.8% of applicants are rejected from all of their choices, in comparison to 2.1% under the three-phase Tunisian mechanism. Virtually all of the students unassigned under the standard mechanism are in the bottom half of the priority distribution. The sequential implementation brings low-priority students closer to updated information about vacancies. This increases their probability to receive their perfect-information matches by preventing them from filling their list with programs that have already filled up. High-priority students’ gains result from a fallback-option channel. Under the standard one-phase restricted-list implementation, a student rejected from all of her listed choices would then be assigned to a program with leftover seats once all the students have gone through the application/assignment process. The cost of rejection is particularly high for high-priority students because the pool of programs with leftover seats at the end is very different from the pool of programs they could get admitted to conditional on applying at their priority level. Because they maximize ex ante expected utility, high-priority students are induced to choose a list with a very low probability of rejection, which in turn leads them to be matched to higher admission-probability, lower-utility programs. The sequential implementation reduces the expected cost of rejection, inducing them to apply to lower admission-probability, higher-utility alternatives.

The large gains from sequencing are partly explained by applicants’ subjective expectations about their admission chances. Estimates from the expectations-formation model show that low-priority students behave as if their admission chances to all programs were higher than they actually are —as if the selectivity of all programs would decrease relative to the previous year. This amplifies their propensity to list programs that, ex post, turn out to be full. Estimates from the expectations-formation model show that, on the contrary, high-priority students tend to behave as if their admission chances to all programs were lower than they actually are —as if the selectivity of programs would increase relative to the previous year. This magnifies their tendency to apply to low-selectivity program in a one-phase setting.

This paper contributes to several strands of the literature. It adds evidence to the small empirical literature on the DA. The theoretical literature on mechanism design is large and influential, and in the context of school choice, it is part of an active dialogue between economists and policy-makers (Abdulkadiroğlu et al., 2006). Despite the widespread use of the DA, there is little empirical evidence of the consequences of a central feature of its implementation —list-size restrictions.2 Ajayi and Sidibé (2017) is, to my knowledge, the only empirical paper that looks at this question. Using data from Ghana, where the restricted-list DA is used to assign students to high schools, they quantify the effect of changing the number of programs to which students are allowed to apply. Fack, Grenet, and He (2019) also document strategic behavior in assignment systems based on the restricted-list DA. In their analysis of the high-school match in Paris, they test and reject the null hypothesis that students are truth-telling. These two recent analyses deliver an empirical counterpart to the experimental findings in Calsamiglia, Haeringer, and Klijn (2010). In contrast, Abdulkadiroğlu, Agarwal, and Pathak (2017) provide empirical evidence from the high-school match in New York City that students may find it optimal to truthfully report their preferences even when constrained to submit a application list strictly smaller than their choice set. My paper is the only one build on quasi-experiment to recover students’ preferences and expectations, and to study the effects of a practical and simple policy that can mitigate the consequences of list-size restrictions by sequentially providing students with updated information about vacancies and high-quality fallback options.

The information-revelation channel and the fallback-option channel documented in this paper are consistent with the findings reported in recent papers documenting students’ application behaviors when they are constrained in their number of applications, and do not know a priori which schools would offer them admission. Pallais (2015) finds that, when allowed to send their ACT scores to a fourth school at no cost, some applicants would send their scores to colleges more selective than those they would have chosen if they could only send their score to three schools for free; while others would would send their scores to colleges less selective. In the lab, Calsamiglia et al. (2010) study students’ application strategies under restricted-list mechanisms, and find that the choices made by subjects with a high valuation for their guaranteed-admission

2A number of studies have compared the unrestricted-list DA to alternative mechanisms (e.g., Agarwal & Somaini, 2018; Calsamiglia, Fu, & Güell, 2014; He, 2016; Kapor et al., 2018).
option are significantly different from the choices made by subjects with a low valuation for their guaranteed option. More recently, Akbarpour and van Dijk (2018) examine theoretically the fallback-option channel, putting an emphasis on its consequences in settings in which fallback options consist in alternatives that are not part of the centralized public school system (e.g., private schools), which not all students equally have access to. My paper quantifies the magnitude of these two channels using a choice model estimated with observational data, and illustrates the potential welfare gains from a simple modification of a very common assignment procedure.3

In this paper, students’ preferences are identified from a subset of truthful applicants to then recover students’ subjective expectations about their admission chances. In this regard, the paper closely relates to studies documenting the effect of providing information or incentives for truthtelling in contexts where students’ subjective expectations may not coincide with their true admission probabilities. Hastings and Weinstein (2009) conduct a field experiment in Charlotte-Mecklenburg and show that providing families with their child’s odds of admission changes application behaviors, which implies that applicants’ expectations about their admissions chances differ from their true odds of admission. More recently, conducting a survey of parents of kindergartners and ninth-graders participating in the New Haven school-choice mechanism, Kapor et al. (2018) show that subjective beliefs about their child’s admission chances differ from true admission probabilities, and that the magnitude of the deviation depends on parental effort and demographics. In a counterfactual exercise, they then compare application and assignments under the Boston mechanism and the strategy-proof DA. My work complements these studies in terms of both methods and setting. I recover subjective expectations from administrative data using a structural model and, I consider college applications under a restricted-list DA, rather than under variance of the Boston mechanism.

The paper is organized as follows. The next section presents the student’s application problem and the sequential assignment mechanism. It reviews standard properties of the DA, describes the inefficiencies list-size restrictions can generate, and illustrates how a sequential implementation can offset them. Section 3 introduces the empirical setting. It presents the college assignment procedure in Tunisia, describes the sample, and provides reduced-form estimates of how the sequential information revelation affects application behaviors. Section 4 presents the empirical model of application portfolio choice. Section 5 describes the two-step identification strategy. Section 6 shows the estimated preferences and expectations parameters. Finally, Section 7 compares students’ outcomes under the sequential DA procedure and the standard implementation of the restricted-list DA, and discusses the value of information and fallback options in a centralized school choice system. Section 8 concludes.

2 The Applicants’ Problem, Information, and Fallback Options

This section sets up the portfolio choice problem that will be brought to the data, reviews basic properties of the DA that will ground the identification strategy, and gives some intuition about the channels through which a sequential implementation of the DA can affect students’ applications and assignments.

2.1 The Applicants’ Problem and a Common Assignment Mechanism

I consider the application problem faced by students in a school choice setting4. A typical centralized school choice mechanism involves (1) students simultaneously submitting an ordered list of academic programs; and (2) a pre-specified rule or algorithm based on which the clearing house processes application lists to assign students to schools.

3 Beyond school choice and college applications, revelation of information and changes in fallback options are channels that can affect choices and outcomes in any situation of decision-making under uncertainty. When it comes to empirical analyses of assignment mechanisms, Waldinger (2017) analyzes the consequences of heterogeneity in outside options among public housing applicants.

4 Although the empirical context considers applications to universities, I use the term school choice because refers to one-sided many-to-one matching problems; while college admissions refer to two-sided many-to-one matching problems. In the context of college admissions, students and schools both have preferences over the other side of the market.
In this paper, I assume that at the time of applying, each applicant $i$ knows the flow utility she would derive from any of the $K$ alternatives in her choice set, and that she has (possibly subjective) expectations about her probability to be admitted to the different programs.\(^5\) She chooses the application list $L_i$ that, among all ordered lists of up to $M$ alternatives, maximizes her expected utility:

$$EU_i(L_i) = \sum_{k=1}^{M} \left[ \pi_i(L_i(k)) \times u_i(L_i(k)) \right] + \bar{\pi}_i \times V_i(0)$$

where $M$ is the maximum number of alternatives students are allowed to rank in their list, $\pi_i(L_i(k))$ denotes $i$’s expectations about her assignment probability to the $k^{th}$-ranked element of her application list; $u_i(L_i(k))$ denotes the flow utility derived from admission to this $k^{th}$-ranked element; $V_i(0)$ denotes the value of being left unassigned; and $\bar{\pi}_i$ denotes student $i$’s probability to be rejected from all her listed choices.

This paper focuses on implementations of the DA algorithm. In the simple case when a unique priority ordering is used by all schools, the DA proceeds as follows:\(^6\)

The serial dictatorship

**Step 1/** The first-ranked student is assigned to her first-listed program.

**Step $(k+1)/** For any $k \geq 1$, once the $k^{th}$ student in the priority ranking has been assigned, the student ranked $(k+1)^{th}$ is assigned to the highest-ranked element of her list that still has a vacancy. If all of her listed choices are full at that point, she is left unassigned and the algorithm proceeds to the next student.

**Stop/** The algorithm stops after all students have been processed.

If students can list as many programs as there exist in their choice set (i.e $M = J$), an assignment mechanism based on the DA algorithm is strategy-proof: it is a dominant strategy for any student to simply list all programs by decreasing order of preference (Dubins & Freedman, 1981; Gale & Shapley, 1962; Roth, 1982). In addition, the resulting match is non-wasteful and free of justified envy: no student prefers to her assignment a school which eventually does not fill up, or which admitted a student with lower priority than her (Abdulkadiroğlu & Sönmez, 2003; Gale & Shapley, 1962). When students can only apply to $M < J$ alternatives, there is no general dominant strategy, the desirable properties of match are not guaranteed. The simple example below illustrates, in this case, the strategic incentives faced by applicants and the ex post inefficiencies that can arise. After that, two propositions give partial characterizations of students’ behavior that will be useful in the rest of the paper.

**A simple example.** Suppose there are two programs A and B, each with two seats. Suppose there are four students, ranked from 1 to 4 by strict priorities. For students $i = S1,...,S4$, and preferences for programs are given by $u_{iA} = \bar{u}_A + \varepsilon_{iA}$ and $u_{iB} = \bar{u}_B + \varepsilon_{iB}$, where $\varepsilon_{iA}$, $\varepsilon_{iB}$ are independent, for instance $i.i.d. N(0,1)$. Students who do not get assigned to any program obtain the outside option, which yields utility 0. For simplicity, suppose $\bar{u}_A$ and $\bar{u}_B$ are large enough so that the probability that a student $i$ prefers the outside option to both schools is negligible. Assume students know their priority ranking, and that there are twice two seats to be apportioned; assume preferences for programs are private information, but that their distribution is common knowledge. In a benchmark perfect-information setting, each student would simply name her most-preferred programs among those still available for her priority level and be assigned to it. Here, I suppose students can only apply to one program and I describe how the restricted-list application strategies and assignment differ from the perfect-information outcomes. Each student solves problem (1). It is clear that S1 and S2 will always “truthfully” list their most preferred program, and be assigned to it.

\(^5\)Throughout the paper, I use the term admission chance or admission probability to designate $i$’s probability to clear the admission cutoff of a program $j$.

\(^6\)This simple case is the one relevant for the empirical analysis in this paper. When a unique priority ordering is used by all schools, the DA boils down to the serial dictatorship algorithm. A more general version of the DA allows for school-specific priorities.
However, S3 and S4 may not report their most-preferred program with respective probabilities:

\[ P(EU_3(\{A\}) < EU_3(\{B\})) \] and \( u_{3A} > u_{3B} \)

and \( P(EU_4(\{A\}) < EU_4(\{B\})) \) and \( u_{4A} > u_{4B} \)

Assignments may differ from the perfect-information assignment for the following three reasons: (i) S3 is unassigned, which happens if either she applies to A and A is full, or she applies to B and B is full. This occurs with probability:

\[
P(u_{1,A} > u_{1,B}) \times P(u_{2,A} > u_{2,B}) \times P(EU_3(\{A\}) > EU_3(\{B\})) + \]

\[
P(u_{1,A} < u_{1,B}) \times P(u_{2,A} < u_{2,B}) \times P(EU_3(\{A\}) < EU_3(\{B\}))
\]

(ii) S3 is ex post-suboptimally assigned, that is, either she applies to A while she prefers B and gets assigned to A while B has seats left; or she applies to B while prefers A and gets assigned to B while A has seats left. This happens with probability:

\[
[P(u_{1,A} > u_{1,B}) \times P(u_{2,A} < u_{2,B}) + P(u_{1,A} < u_{1,B}) \times P(u_{2,A} > u_{2,B})] \times
\]

\[
[P(EU_3(\{A\}) > EU_3(\{B\}) \text{ and } u_{3,B} > u_{3,A}) + P(EU_3(\{A\}) < EU_3(\{B\}) \text{ and } u_{3,B} < u_{3,A})]
\]

(iii) S4 is unassigned, which occurs with probability:

\[
P(EU_4(\{A\}) > EU_4(\{B\})) \times
\]

\[
[P(u_{1,A} > u_{1,B}) \times P(u_{2,A} > u_{2,B}) + P(u_{1,A} > u_{1,B}) \times P(u_{2,A} < u_{2,B})] \times
\]

\[
(P(u_{1,A} > u_{1,B}) \times P(u_{2,A} < u_{2,B}) + P(u_{1,A} < u_{1,B}) \times P(u_{2,A} > u_{2,B}))
\]

Loosely speaking, (i) and (iii) can be seen as cases of overshooting: the students are unassigned because they applied to programs that end being too selective for their respective priority levels. (ii) can be seen as a case of undershooting: the student ends up assigned to a relatively low-utility program, while she would have been assigned to a higher-utility program if she had included it in her application list.

Although it may be dominant to report a vector of programs that differs from one’s most-preferred vector, Proposition 1 (Haeringer & Klijn, 2009) establishes that one never benefits from ranking the reported alternatives differently than by decreasing order of preference. Because this will be useful for the empirical analysis, Proposition 2 establishes a sufficient condition for truth-telling to be a dominant strategy.\(^7\)

**Proposition 1.** [Haeringer and Klijn (2009)] (a) If a student finds at most M schools acceptable, then she can do no better than submitting her true preferences.

(b) If a student finds more than M schools acceptable, then she can do no better than employing a strategy that selects M schools among the acceptable schools and ranking them according to her true preferences.

**Proposition 2.** (a) Condition 1 (below) is a sufficient condition for students not to have a strict incentive to misreport their preferences over their choice set.

(b) Under Assumption 1, Condition 1 is a sufficient condition for students not to misreport their preferences over their choice set.

**Condition 1.** Student \( i \) thinks that with probability 1, (at least) one of her \( M \) most-preferred programs will have a seat left for her.

**Assumption 1.** When indifferent between doing so or not, a student does not mis-represent her unconstrained preference ranking. In other words, a student does not report her most-preferred programs in her application list only when it is strictly profitable to do so.

\(^7\)A proof is given in Online Appendix B.
2.2 Sequential Implementation

Instead of the standard (one-phase) implementation of the DA, which involves all $N$ students simultaneously submitting their application lists and then being assigned, this paper considers a sequential implementation. It involves first dividing the cohort in $K \leq N$ assignment groups that will both submit lists and be assigned, one after the other. In the case in which the same priority order is used by all schools, the division of the cohort can be straightforwardly made along this priority order. Suppose the $N$ students are ranked by a strict priority order from 1 to $N$. Let $K_1, K_2, \ldots, K_K$ be the sizes of each of the $K$ groups to be created. Assign students with priority ranks 1 to $K_1$ to Group 1, students with priority ranks $K_1 + 1$ to $K_1 + K_2$ to Group 2, etc. Priority order is preserved within groups. Given these groups, the assignment procedure goes as follows:

**$K$-phase DA**

*Phase 1/* The number of seats open in each program is publicly revealed. Group-1 students submit application lists, and are then assigned using the DA algorithm.

*Phase $(n+1)/ For any $1 \leq n \leq (K-1)$, vacancies remaining after the assignment of Group-$n$ students are publicly revealed. Unassigned Group-$n$ students are added at the top of Group-$(n+1)$ according to their initial priority order. Group-(n+1) students submit application lists, and are then assigned using the DA algorithm.

In the limit, if $K$ is equal to the number of students, the sequential implementation guarantees that each student can choose perfect information, and simply needs to reveal their most-preferred program among those he knows are available in order to be assigned to it.

**The simple example, continued: information revelation and fallback option.** (i.) Suppose the application/assignment process is sequenced with the following two groups: $\{1, 2\}, \{3, 4\}$. Then, S3 has perfect information so maximizing (1) leads her to apply and get assigned to her most-preferred alternative among those available. If the information revealed to S3 is that both seats in A or both seats in B are available, then S4 also chooses under perfect information and gets assigned to her most-preferred programs among those available. However, with probability:

$$1 - \left[ P(u_{1,A} > u_{1,B}) \times P(u_{2,A} > u_{2,B}) + P(u_{1,A} < u_{1,B}) \times P(u_{2,A} < u_{2,B}) \right],$$

the information revealed to Group 2 is that each school has one seat. In that case, S4 will not choose under perfect information. She may not apply to her most-preferred program, and she may be left unassigned.

(ii.) Suppose instead that the application/assignment process is sequenced as follows: $\{1, 2, 3\}, \{4\}$, and we tell students 1, 2, and 3 that they will be pooled at the top of the next application group if they fail to be assigned. Then, both S3 and S4 are always assigned to their most-preferred programs among those available. In the first phase, S3 compares the expected utility of listing A vs B, taking into account the new fallback option given to her in case she fails to be assigned:

$$Eu_A = P(\{A \text{ has a seat left}\}) \times u_A + (1 - P(\{A \text{ has a seat left}\})) \times u_B$$

$$Eu_B = P(\{B \text{ has a seat left}\}) \times u_B + (1 - P(\{B \text{ has a seat left}\})) \times u_A$$

Then, S3 always applies to her most-preferred program. If the program she applies to has a seat available, S3 is assigned to it, and the procedure moved to the next application phase, revealing the updated choice set. S4 is the first student in line to be assigned in the second phase, and chooses under perfect information. There is one seat to choose, and she gets assigned to it. If, however, S3’s most-preferred program (say, A) was filled by S1 and S2, S3 is rejected from her listed choice and she is pooled at the top of the next application group. In the second phase, it is revealed that the other program (here, B) has two seats available, so both S3 and S4 apply under perfect information and are assigned the their most-preferred program among those
available (here, B).

The goal of this paper is to evaluate the extent to which, in practice, a sequential implementation of the DA can result in higher-utility matches than the standard single-phase procedure. The empirical analysis is based on university application data from Tunisia, where a three-phase DA is used.

3 The University Match in Tunisia

This section introduces the empirical setting of this paper. Taking advantage of the cutoffs generated by the division of the applicant pool into three groups, I provide reduced-form evidence of the consequences of the sequential implementation on students’ application behaviors and assignments.

3.1 Institutional Background and Sample Description

Institutional background. Every June, high-school seniors in Tunisia take the national end-of-high-school exam. In high school, students choose one major (section) out of six possible, and conditional on this major, every high-school senior in the country takes the same exam at the same time. Passing this exam (that is, scoring at least 10 out of 20 on average over the eight to ten tests of the exam) is a sufficient and necessary condition to graduate from high school and gain access to public post-secondary education in Tunisia. Universities deliver field-specific degrees that require the completion of a standard curriculum approved by the Ministry of Higher Education. Undergraduate students typically start specializing in one field of study from their first semester. In this paper, I refer to a pair (university, field) as a program.

Assignment of high-school graduates to post-secondary programs is determined by a nationwide mechanism that is virtually similar to the three-phase serial dictatorship described in Section 2.

On paper, the assignment procedure simultaneously involves students from all high-school majors, and uses, rather than the serial dictatorship, a DA algorithm with a few distinct program-specific priorities. However, in practice, the situation faced by students graduating from high school with a Math major (henceforth, math-majors) turns out to be very similar to a three-phase serial dictatorship on a market independent from other high-school graduates. Students’ grades at the national end-of-high-school exam are used to determine a baseline priority score. Within each high-school major, students are ranked according to this baseline priority score, and, for each major, the cohort is divided into three groups—the top 30% within the major, middle 40%, and bottom 30%. A very large majority of programs, which all together have enough seats accommodate all math-majors, allocate their seats based on the baseline priority score and have a designated quota of seats for math-majors. Only 5% of the seats math-majors can apply to are allocated using a different priority score. All these seats are in languages programs and the correlation between the alternative priority score (a weighted sum of the baseline priority score and grades at the national language test) and the baseline priority score is close to 1. Finally, programs that open their seats to all high-school majors in specific fields math-majors do not have a predilection for (e.g., Arts, Languages) and they represent only about 3% of applications and 2% of assignments of math-majors. Online Appendix C describes the actual procedure in detail, gives statistics about the programs considering cross-major applicants or using an alternative priority, and argues further that the empirical simplification is legitimate. In the rest of this paper, I treat the data as if the serial dictatorship was used to assign students, and as if math-majors applied on a market independent from other high-school graduates.

I now summarize the implementation features that will be exploited in the empirical analysis. In year 2010, and for the sample of math-majors, the common priority ranking of students used by programs to determine applicants’ admission is based on a priority score, which can be viewed as a standardized test score and is known to the student. This score is determined as a weighted average of the student’s

---

8The empirical analysis here focuses on Math-majors. The five other majors are: Literature, Experimental sciences, Economics/Management, Computer Science, and Technology.

9See qualification in the previous paragraph and details in Online Appendix C.
grades at the various tests of the national end-of-high-school exam.\textsuperscript{10} Students may submit an ordered list of up to 10 post-secondary programs. Before submitting applications, all high-school graduates are given a handout containing information about the available post-secondary programs over the country. The handout indicates, for each existing program, the number of vacancies open for the next academic year and the past-year admission cutoff, that is, the priority score of the marginal student admitted in the previous year. The application/assignment process is split into three successive phases. Applicants are ranked in decreasing order of priority, and the cohort is divided into three groups based on this priority ranking —the top 30\% of students (“Group 1” students), the middle 40\% (“Group 2”), and the bottom 30\% (“Group 3”). After each group has gone through the assignment process, the number of vacancies in each program is publicly updated, so next-group students are told which vacancies remain before submitting their application list. Students who fail to be admitted to any of their listed choices are pooled at the top of the next application group (if there is one) and proceed to submitting a new application list after the information about vacancies is publicly updated.\textsuperscript{11} If there is no next application group, unmatched students are administratively assigned to leftover seats.\textsuperscript{12}

Table 1: Descriptive statistics: students

<table>
<thead>
<tr>
<th>Demographics</th>
<th>All</th>
<th></th>
<th>Group 1</th>
<th></th>
<th>Group 2</th>
<th></th>
<th>Group 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.dev.</td>
<td>Mean</td>
<td>S.dev.</td>
<td>Mean</td>
<td>S.dev.</td>
<td>Mean</td>
<td>S.dev.</td>
</tr>
<tr>
<td>Female</td>
<td>.53</td>
<td>.50</td>
<td>.52</td>
<td>.50</td>
<td>.54</td>
<td>.50</td>
<td>.52</td>
<td>.50</td>
</tr>
<tr>
<td>High SES</td>
<td>.60</td>
<td>.49</td>
<td>.78</td>
<td>.41</td>
<td>.58</td>
<td>.49</td>
<td>.47</td>
<td>.50</td>
</tr>
<tr>
<td>From Tunis</td>
<td>.30</td>
<td>.46</td>
<td>.33</td>
<td>.47</td>
<td>.30</td>
<td>.46</td>
<td>.27</td>
<td>.44</td>
</tr>
<tr>
<td>From Coast (excl. Tunis)</td>
<td>.48</td>
<td>.50</td>
<td>.53</td>
<td>.50</td>
<td>.49</td>
<td>.50</td>
<td>.43</td>
<td>.49</td>
</tr>
<tr>
<td>From South</td>
<td>.03</td>
<td>.17</td>
<td>.01</td>
<td>.11</td>
<td>.03</td>
<td>.17</td>
<td>.05</td>
<td>.22</td>
</tr>
<tr>
<td>Priority and academic perf.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stdized priority score</td>
<td>0</td>
<td>1</td>
<td>1.28</td>
<td>.46</td>
<td>-.13</td>
<td>.38</td>
<td>-1.10</td>
<td>.34</td>
</tr>
<tr>
<td>STEM high-school</td>
<td>0</td>
<td>.85</td>
<td>1.04</td>
<td>.39</td>
<td>-.09</td>
<td>.40</td>
<td>-.92</td>
<td>.33</td>
</tr>
<tr>
<td>non-STEM high-sch. perf.</td>
<td>0</td>
<td>.79</td>
<td>.75</td>
<td>.54</td>
<td>-.07</td>
<td>.58</td>
<td>-.66</td>
<td>.59</td>
</tr>
<tr>
<td>Applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>List 10 choices</td>
<td>.70</td>
<td>.46</td>
<td>.67</td>
<td>.47</td>
<td>.76</td>
<td>.43</td>
<td>.65</td>
<td>.48</td>
</tr>
<tr>
<td>Number of choices listed</td>
<td>9.02</td>
<td>1.8</td>
<td>8.86</td>
<td>1.9</td>
<td>9.3</td>
<td>1.5</td>
<td>8.86</td>
<td>1.9</td>
</tr>
<tr>
<td>List all programs in STEM</td>
<td>.36</td>
<td>.48</td>
<td>.19</td>
<td>.40</td>
<td>.46</td>
<td>.50</td>
<td>.40</td>
<td>.49</td>
</tr>
<tr>
<td>List all prog. in same region</td>
<td>.26</td>
<td>.44</td>
<td>.10</td>
<td>.31</td>
<td>.35</td>
<td>.48</td>
<td>.31</td>
<td>.46</td>
</tr>
<tr>
<td>Assignments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admitted to 1st listed prog.</td>
<td>.39</td>
<td>.49</td>
<td>.45</td>
<td>.50</td>
<td>.39</td>
<td>.49</td>
<td>.35</td>
<td>.48</td>
</tr>
<tr>
<td>Admitted to 2nd listed prog.</td>
<td>.15</td>
<td>.36</td>
<td>.16</td>
<td>.36</td>
<td>.16</td>
<td>.37</td>
<td>.12</td>
<td>.33</td>
</tr>
<tr>
<td>Admitted to 3rd listed prog.</td>
<td>.10</td>
<td>.30</td>
<td>.11</td>
<td>.31</td>
<td>.11</td>
<td>.31</td>
<td>.07</td>
<td>.26</td>
</tr>
<tr>
<td>Admitted to 4th listed prog.</td>
<td>.07</td>
<td>.26</td>
<td>.08</td>
<td>.28</td>
<td>.07</td>
<td>.26</td>
<td>.06</td>
<td>.24</td>
</tr>
<tr>
<td>Admitted to 5th listed prog.</td>
<td>.05</td>
<td>.22</td>
<td>.06</td>
<td>.23</td>
<td>.05</td>
<td>.22</td>
<td>.05</td>
<td>.21</td>
</tr>
<tr>
<td>Admitted to 6th listed prog.</td>
<td>.04</td>
<td>.19</td>
<td>.05</td>
<td>.20</td>
<td>.04</td>
<td>.20</td>
<td>.03</td>
<td>.18</td>
</tr>
<tr>
<td>Admitted to 7th+ list. prog.</td>
<td>.06</td>
<td>.12</td>
<td>.02</td>
<td>.07</td>
<td>.07</td>
<td>.13</td>
<td>.11</td>
<td>.16</td>
</tr>
<tr>
<td>Administratively assigned</td>
<td>.02</td>
<td>.14</td>
<td>0</td>
<td>.06</td>
<td>0</td>
<td>.07</td>
<td>.06</td>
<td>.23</td>
</tr>
<tr>
<td>Admitted in later round</td>
<td>.02</td>
<td>.14</td>
<td>.02</td>
<td>.13</td>
<td>.04</td>
<td>.19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample size</td>
<td>10,884</td>
<td>3,287</td>
<td>4,355</td>
<td>3,242</td>
<td>10,884</td>
<td>3,287</td>
<td>4,355</td>
<td>3,242</td>
</tr>
</tbody>
</table>

Note: In the second panel, STEM (resp. non-STEM) high-school performance is the unweighted average of the student’s standardized scores at the Math, Physics, Natural Sciences, and Comp. Sci. (resp. English, French, Arabic, and Philosophy) tests of the end-of-high-school national exam.

\textsuperscript{10} The priority score is observed in the data. A number of end-of-high-school exam grades are also observed.

\textsuperscript{11} The new list is formed based on the programs available at the time it is submitted, and not based on the programs available when the student submitted her initial list. In the data, only the very last list submitted by each student is recorded.

\textsuperscript{12} I do not have information about the rule used to determine administrative assignments. Considering that geographical location is an important criterion used by the administrator, in the counterfactual exercise, I will assume that, at this stage, a student is randomly assigned to one of the 25\% of leftover seats that are closest to her home region.
**Students.** I focus my analysis on the 11,029 students graduating high school with the Math major. Among them, I drop 145 students who either did not submit an application list, or for whom high-school information and/or all end-of-high-school test scores are missing. Online Appendix C details how I define the high-SES binary indicator; how I construct measures of high-school performance and of distance between students’ home region and programs locations; and how I treat missing demographic information. Table 1 describes the 10,884 students in the final sample, as well as the division of the cohort into the three application groups. The top panel of Table 1 shows the distribution of demographics that typically affect schooling decisions. About half of the sample is female, 60% have a high-SES background, about 80% are from Tunis and the coastal regions. While the sex ratio is constant across groups, the share of high-SES students and students from the most economically dynamic regions of the country decreases as priority (that is, high-school performance) goes down. A number of application behaviors and assignment patterns are similar across priority groups, while others differ. Overall, students list on average 9 choices out of the 10 possible, and 70% of them actually list 10 choices. More than 40% of Group-2 and Group-3 students list only programs in STEM fields, against 20% of Group-1 students. More than 30% of Group-2 and Group-3 students list all of their choices in the same geographic region, against only 10% of Group-1 students. While high- and low-SES students are not identically distributed along the priority ranking, they have similar application behaviors conditional on priority ranking (see for instance Figure A.1 in Online Appendix A). Conditional on being assigned to one of their choices, about 45% of Group-1 students get assigned to their first-listed choice, against 39% and 35% of Group-2 and Group-3 students. 80% of Group-1 students and 73% of Group-2 students are assigned to one of their top four listed choices, against only 60% of Group-3 students. 6% of Group-3 students end up administratively assigned, while few Group-1 and Group-2 students get rejected from all their listed choices.

Table 2: Descriptive statistics: programs

<table>
<thead>
<tr>
<th></th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice set size</td>
<td>616</td>
<td>562</td>
<td>290</td>
</tr>
<tr>
<td><strong>Fields</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non STEM</td>
<td>.33</td>
<td>.34</td>
<td>.40</td>
</tr>
<tr>
<td>% Humanities</td>
<td>.04</td>
<td>.05</td>
<td>.04</td>
</tr>
<tr>
<td>% Arts</td>
<td>.11</td>
<td>.13</td>
<td>.17</td>
</tr>
<tr>
<td>% Tertiary</td>
<td>.05</td>
<td>.01</td>
<td>.00</td>
</tr>
<tr>
<td>% Econ/ Business/ Mgmt</td>
<td>.11</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>% Social Sciences/ Law</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>% STEM</td>
<td>.67</td>
<td>.66</td>
<td>.60</td>
</tr>
<tr>
<td>% Health/Life Sciences</td>
<td>.09</td>
<td>.05</td>
<td>.02</td>
</tr>
<tr>
<td>% Earth Sciences</td>
<td>.03</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>% Math/ Comp. Sci.</td>
<td>.11</td>
<td>.16</td>
<td>.20</td>
</tr>
<tr>
<td>% Physics/ Chem./ Engin.</td>
<td>.44</td>
<td>.41</td>
<td>.34</td>
</tr>
<tr>
<td><strong>Degrees</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% BA</td>
<td>.36</td>
<td>.48</td>
<td>.59</td>
</tr>
<tr>
<td>% BS</td>
<td>.29</td>
<td>.33</td>
<td>.39</td>
</tr>
<tr>
<td>% Advanced</td>
<td>.34</td>
<td>.19</td>
<td>.03</td>
</tr>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% in Tunis</td>
<td>.30</td>
<td>.21</td>
<td>.07</td>
</tr>
<tr>
<td>% on coast</td>
<td>.52</td>
<td>.58</td>
<td>.56</td>
</tr>
<tr>
<td>% abroad</td>
<td>.01</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Note: percentages given out of the total number of available seats (not programs).*

**Choice set.** The first column of Table 2 shows that, in 2010, 616 programs have seats open for students graduating from high-school with a Math major. Programs are offered in nine different fields of study.\(^1\) Each post-secondary program is identified in the data a six-digit code that reflects the classification of fields used by the Ministry of Higher Education. I use this classification too. The 'tertiary' field consists of programs in specialized education, medicine, and law.

---

\(^1\) Students have high school information and/or all end-of-high-school test scores missing. 65 students are recorded in the data as not having submitted any application list — 3 are Group 1, 18 Group 2, and 44 Group 3. In addition, I drop 55 students, whose application list comprises only programs out of their choice set and/or some of the six programs I drop because they do not have an equivalent in the previous year. See Online Appendix C.
four of which are in STEM. 67% of the seats initially offered to the applicant sample are in STEM fields. Two thirds of seats are offered in programs preparing for the equivalent of a BA or a BS, the last third prepares for advanced degrees.\footnote{I use Bachelor of Arts (BA) and Bachelor of Science (BS) as rough equivalents for Licence appliquée (LA) and Licence fondamentale (LF), respectively, both of which are three-year degrees delivered by universities. Advanced-degree programs include: (i) programs with a curriculum of six years or more (cursus long), that is, delivering degrees in medicine, dental medicine, pharmacy, and architecture; and (ii) classes préparatoires, which are two/three-year programs which purpose is to prepare students to later take competitive exams to enter elite engineering and scientific research schools.} About 80% of seats are offered by universities located in Tunis and on the coast. As part of the sequential implementation, students in Groups 2 and 3 are told which seats were filled by previous-group students; hence, their actual choice set is only a subset of the initial set of all programs, as illustrated by the last two columns of Table 2. By the end of the first application phase, 54 programs have filled up; by the end of the second phase, 326 programs are no longer available. By the time Group-3 applicants submit their applications, most of the seats in Tunis and virtually all the seats in programs preparing to advanced degree have been claimed; very few seats in Health and Life Sciences or Economics/ Business/ Management are left; while most of the seats in Arts are still available. This variation in the choice set induced by the sequential implementation implies that, as some combinations of fields of study, locations, and type of degrees become unavailable, students need to apply to and express preferences over alternatives with other combinations of characteristics. This will help the identification of students’ preferences for discrete program characteristics.

**Cutoffs.** The left-hand-side panel of Figure 1 shows, for each program in the 2010 choice set, its 2010 cutoff against its 2009 cutoff. Two patterns are noticeable. First, there is a strong positive correlation between cutoffs from one year to the next, suggesting that previous-year cutoffs are informative about the realization to come. Second, while there is little year-to-year variation in the cutoffs of the most selective programs, the cutoffs of less selective programs are quite volatile. This is also illustrated by the right panel of Figure 1. The three thick gray lines show the empirical probability density functions (henceforth PDFs) of the difference \( \text{cutoff}_{j,2010} - \text{cutoff}_{j,2009} \) for three sets of programs: all programs (solid black line), the 50\% of programs with lowest cutoffs in 2009 (dotted gray line), and the 50\% of programs with highest cutoffs in 2009 (dashed lightest gray line). To exemplify the increasing variance in cutoffs as selectivity of programs decreases, the thinnest lines show the PDFs of three normal distributions that are close to these empirical PDFs. Pooling all programs, the empirical PDF of cutoff changes is well approximated by a normal with mean 0 and standard deviation .41. By comparison, the empirical PDF of cutoff changes for the least (resp. most) selective programs is well approximated by a normal with mean 0 and standard deviation .58 (resp. .35). This points to the fact that students lower in the priority ranking face a higher level of uncertainty than students early in the priority ranking.

### 3.2 Local Effects of the Sequential Implementation on Application Behaviors and Assignments

Figure 2 illustrates the changes in behaviors observed at the group cutoffs. It plots, as a function of students’ priority, the rank in their application list of the choice they are assigned to. Top-ranked students, at the left, are assigned to their first-listed choice. As priority goes down in Group 1, and as popular programs fill up, students get assigned to increasingly lower choices in their application portfolio. When updated information about vacancies is provided, at the limit between Groups 1 and 2, application behaviors change in such a way that students get assigned to their top-listed choice again. The cycle starts again until the next revelation of information, at the limit between Groups 2 and 3. In this subsection, I further document local application and assignment changes at the information revelation cutoffs.

The division of the applicant pool into application groups creates cutoffs that I use in a sharp regression discontinuity (RD) design in order to document the local effects of providing updated information to applicants.\footnote{Standard graphical evidence supporting the sharpness and validity of the RD design can be found in Online Appendix A.} The local effect of a change in groups on any outcome \( Y \) of interest is estimated by local linear physical education, and services of the tourism, hotel and restaurant industries.
regression (Imbens & Lemieux, 2007):

$$\min_{\alpha, \beta, \tau, \gamma} \sum_{i=1}^{N} 1_{[-h \leq T_i \leq c + h]} \cdot \left( Y_i - \left[ \alpha + \beta (T_i - c) + \Delta 1_{[T_i < c]} + \gamma (T_i - c) 1_{[T_i < c]} \right] \right)$$

where $T_i$ is student $i$’s priority score (running variable), $c$ denotes the group cutoff, and $h$ is the estimation bandwidth. $1_{[T_i < c]}$ is a indicator of $i$ being assigned to the ‘informed group’ (that is, Group 2 at the Group 1/ Group 2 cutoff; and Group 3 at the Group 2/ Group 3 cutoff) or not, $(T_i - c)$ is the distance of $i$’s score to the group cutoff, and $(T_i - c) 1_{[T_i < c]}$ is an interaction term that allows the slope of distance to the cutoff to differ on either side of the group cutoff. $\Delta$ is the coefficient of interest; it measures the change outcome $Y$ induced by the revelation of information. I estimate (2) using, as $Y$, various application list characteristics (e.g. length, selectivity measures), and assignment outcomes (e.g. probability of assignment, rank of the choice assigned). In addition, to understand further how changes in application rates at the cutoffs correlate with the information revealed about programs, I estimate, for each program $j$, the change in application rate at the group cutoffs, using the binary indicator of whether or not student $i$ ranked the program in her list as the dependent variable $Y_i^{(j)}$ in equation (2). I then regress the estimated change in application rate on various program characteristics. Results shown in Tables 3 and 4 point to four interesting patterns.

For each outcome and subsample, the ‘optimal’ bandwidth is chosen using the Imbens and Kalyanaraman (2011) method and may vary from one outcome to another.
Table 3: Reduced-form effects of informational updates on application behaviors and assignment patterns

<table>
<thead>
<tr>
<th>Application behaviors</th>
<th>Groups-1/2 cutoff</th>
<th>Groups-2/3 cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Change</td>
<td>Base level</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td></td>
</tr>
<tr>
<td># listed choices</td>
<td>-0.264</td>
<td>9.081</td>
</tr>
<tr>
<td></td>
<td>(0.256)</td>
<td>(1.754)</td>
</tr>
<tr>
<td>Obs.</td>
<td>727</td>
<td></td>
</tr>
<tr>
<td>Std. PY cutoff of most selective choice</td>
<td>-0.389***</td>
<td>1.239</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.516)</td>
</tr>
<tr>
<td>Obs.</td>
<td>702</td>
<td></td>
</tr>
<tr>
<td>Std. PY cutoff of least selective choice</td>
<td>-0.413***</td>
<td>-0.301</td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.524)</td>
</tr>
<tr>
<td>Obs.</td>
<td>347</td>
<td></td>
</tr>
<tr>
<td>Avg. std. PY cutoff over listed choices</td>
<td>-0.314***</td>
<td>0.481</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.344)</td>
</tr>
<tr>
<td>Obs.</td>
<td>487</td>
<td></td>
</tr>
<tr>
<td>Pctl-terms PY cutoff of most selective choice</td>
<td>-0.094***</td>
<td>0.845</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Obs.</td>
<td>535</td>
<td></td>
</tr>
<tr>
<td>Pctl-terms PY cutoff of least selective choice</td>
<td>-0.073***</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.180)</td>
</tr>
<tr>
<td>Obs.</td>
<td>819</td>
<td></td>
</tr>
<tr>
<td>Avg. pctl-terms PY cutoff over listed ch.</td>
<td>-0.092***</td>
<td>0.653</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.099)</td>
</tr>
<tr>
<td>Obs.</td>
<td>478</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Assignment patterns</th>
<th>Groups-1/2 cutoff</th>
<th>Groups-2/3 cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability to be assigned</td>
<td>0.025</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.195)</td>
</tr>
<tr>
<td>Obs.</td>
<td>676</td>
<td></td>
</tr>
<tr>
<td># of listed choices eligible to</td>
<td>2.784***</td>
<td>5.406</td>
</tr>
<tr>
<td></td>
<td>(0.318)</td>
<td>(2.397)</td>
</tr>
<tr>
<td>Obs.</td>
<td>767</td>
<td></td>
</tr>
<tr>
<td>% of listed choices eligible to</td>
<td>0.322***</td>
<td>0.589</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.236)</td>
</tr>
<tr>
<td>Obs.</td>
<td>487</td>
<td></td>
</tr>
<tr>
<td>Rank of choice assigned</td>
<td>-1.888***</td>
<td>2.200</td>
</tr>
<tr>
<td></td>
<td>(0.365)</td>
<td>(2.287)</td>
</tr>
<tr>
<td>Obs.</td>
<td>397</td>
<td></td>
</tr>
</tbody>
</table>

The Change column gives the estimated average change in outcome at the group cutoff. Std. errors are reported in parentheses below estimates. Estimation bandwidth is 1/8 of Imbens and Kalyanaraman (2011) optimal bandwidth. * p < 0.05, ** p < 0.01, *** p < 0.001. The Base level column gives control-group statistics about the outcome. Std. deviations are reported in parentheses below mean values. Std. PY cutoff stands for ‘standardized past-year cutoff’, which corresponds to the standardized priority score of the marginally admitted student in 2009. Pctl-terms PY cutoff stands for ‘past-year cutoff in percentiles-terms’, which corresponds to the percentile ranking (within the math-major 2009 cohort) of the marginally admitted student in 2009. Online Appendix C gives more details about the definition of these measures.

Read: The most selective program listed by marginally uninformed students at the Group-1/2 cutoff has, on average, a standardized past-year cutoff of 1.2, which corresponds to the 85th percentile of the priority score distribution. The most selective program listed by their marginally informed counterpart has a standardized past-year cutoff that is .4 std. dev. lower, which corresponds to a decrease of 9 percentiles in the priority score distribution.

(1) Marginally informed students submit shorter and less-selective application lists than marginally uninformed students. The top panel of Table 3 shows that, as compared to students that are marginally non-informed, marginally informed students list slightly fewer choices —.26 fewer at the Group 1/ Group 2 cutoff, and .63 fewer at the Group 2/ Group 3 cutoff, with only the latter difference being statistically different from 0 at conventional levels. Marginally informed students apply to programs that are less selective than their marginally uninformed counterparts. For instance, the most selective of their choices,
has a past-year cutoff that is about .4 standard deviation lower, which corresponds to 9 percentiles of the priority distribution at the Group 1/Group 2 cutoff, and 11 percentiles at the Group 2/Group 3 cutoff. Interestingly, the same is true students’ least selective choices. The least-selective choice listed by marginally informed students has a past-year cutoff that is about .7 percentiles lower in the priority distribution than that of their marginally informed counterparts.

(2) Marginally informed students increase their application rate to less selective and more popular programs among those with remaining vacancies. The top panel in Table 4 shows that a program being declared full (for the first time) at the group cutoff is correlated with a drop in application rate —by 14 and 5 percentage points at the Group 1/Group 2 and Group 2/Group 3 cutoffs, respectively. The middle and bottom panels show that, for programs that are declared full, the magnitude of the drop in application rates increases with the program’s initial number of vacancies, and its selectivity level. Symmetrically, for programs that are declared not to be full, a larger number of remaining vacancies and a higher past-year cutoff are correlated with a larger surge in application rates.

(3) Marginally informed students are more likely to be assigned to an element of their list than their marginally uninformed counterparts. The bottom panel of Table 3 shows that students’ probability to be assigned to one of their listed choices, rather than being rejected from all of them, is increased by 9 percentage points at the Group 2/Group 3 cutoff, and by 2 percentage points at the Group 1/Group 2 cutoff (but this latter effect is not statistically different from 0).

(4) Marginally informed students are assigned to higher-ranked elements of their lists than marginally uninformed students. The bottom panel of Table 3 also shows that marginally informed students end up clearing the ex post admission cutoff of a larger share of their listed choices (+32% at the Group 1/Group 2 cutoff, and +58% at the Group 2/Group 3 cutoff) than marginally uninformed students do. This induces them to be assigned to a higher-ranked element of their list —1.9 and 2.4 ranks higher at the Group 1/Group 2 and Group 2/Group 3 cutoffs, respectively.

Validity of the RD design means that the allocation of students in one group or the next is, locally, as good as random: students on either side of a group cutoff have, on average, the same observable and unobservable characteristics, as well as the same preferences for post-secondary programs. Discontinuities in application behaviors at the group cutoffs are evidence that students do not have perfect foresight and that they use the information they are provided. In addition, Table 4 shows that the changes in application rates are consistent with students understanding and using the information they are given at the group cutoffs.

However, this reduced-form analysis does not inform whether students benefit, in terms of ex post utility, from the sequencing of the assignment procedure. It does not inform either about the behavior and gains of students located further away from the group cutoffs in the priority ranking. Conducting a welfare evaluation of the effects of sequencing requires comparing how students fare in terms of ex post utility under alternative sequencing scenarios. Simulating students’ applications and assignments under alternative scenarios requires to know students’ preferences for programs, and to understand how students derive expectations about their admission chances from the available information. The next section presents the empirical model that I will use to recover preferences and expectations.
Table 4: Correlations between local change in application rates and information

<table>
<thead>
<tr>
<th>Regression</th>
<th>Groups 1/2 cutoff</th>
<th>Groups 2/3 cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Just full</td>
<td>-1.400***</td>
<td>-0.492***</td>
</tr>
<tr>
<td>Constant</td>
<td>0.122***</td>
<td>0.204***</td>
</tr>
<tr>
<td>R-sqr</td>
<td>0.208</td>
<td>0.257</td>
</tr>
<tr>
<td>Obs.</td>
<td>616</td>
<td>616</td>
</tr>
</tbody>
</table>

For all regressions, the outcome variable is ‘Estimate change in application rate at the group cutoff.’ Just full is a indicator equal to 1 if the program was declared to be full at the most-recent group change; Not just full is 1 when Just full is 0. Pctl-terms PY cutoff stands for ‘past-year cutoff in percentiles-terms’ (see Table 3). Bootstrap std. errors in parentheses, account for two-step estimation. * p < 0.05, ** p < 0.01, *** p < 0.001

4 Empirical model

Students choose the application list that maximizes their expected utility (1).

4.1 Utility specification

I assume that student $i$’s utility from being assigned to program $\ell$ depends on the characteristics of program $\ell$ and student $i$’s own characteristics:

$$ u_i(\ell) = \sum_k (\beta_k + \gamma_k \cdot X_i + \alpha_k^{(i)}) \cdot z_{\ell,k} + \sum_m (\beta_m + \gamma_m \cdot X_i + \alpha_m^{(i)}) \cdot w_{i\ell,m} + \varepsilon_{i\ell} $$

where $z_{\ell,k}$ are observable program-specific characteristics, such as field of study, degree prepared, or past-year admission cutoff; $w_{i\ell,m}$ are characteristics that vary both with program and with student, such as distance between the student’s home and the university; and $X_i$ is a vector of student characteristics. The parameter vector $\beta$ captures systematic tastes for program characteristics $z$ and $w$. Parameters $\gamma$ allow tastes for program characteristics to vary systematically with students’ observable characteristics, such as high-school performance or gender. The student-specific parameters $\alpha^{(i)}$ capture unobserved heterogeneity in tastes for program characteristics. Finally, $\varepsilon_{i\ell}$ is an idiosyncratic part of $i$’s utility for $\ell$ that is observed by $i$ but not by the econometrician. I assume the distribution of $\varepsilon_{i\ell}$ to be known (in the estimation, $\varepsilon_i$ are i.i.d. type-1 extreme value) and independent of programs’ and students’ characteristics.

A key restriction I impose is that unobservables $\varepsilon_i$ and $\alpha^{(i)}$ are independent of students’ observable characteristics, in particular distance between students and programs. This rules out, for instance, families systematically choosing their geographical residence at the time of high school to be next to the university
programs students like. This guarantees that coefficients on school attributes identify the students’ valuation for these attribute and does not capture correlated variation with unobservable tastes. Note that programs’ and students’ observable characteristics will be taken as given and fixed in the counterfactual analysis and welfare evaluations.

While $\varepsilon_{it}$ are assumed to be i.i.d. across programs, the random coefficients $\alpha$ allow for within-student correlation in unobserved preferences for program characteristics. Identification of these parameters comes from directly observing substitution patterns through the application list of truthful students. In the empirical exercise, I allow unobserved heterogeneity in tastes for the different STEM fields of study\(^{18}\) as well as for distance, and I further assume:

$$a_{k'}^{(i)} = \nu_{k'} \times a_{k'}^{(i,0)} \quad \text{where} \quad a_{k'}^{(i,0)} \sim i.i.d \ N(0,1)$$

For identification, I normalize to zero the coefficients on a reference field and on a reference terminal degree. This means that, for each student, the value of every post-secondary program is interpreted as relative to the mean value of a local (in that distance traveled is 0) program that is not selective (past-year cutoff is 0 for programs that did not fill to capacity in 2009\(^{19}\)), and upon completion of which the student would earn a BA (the reference degree) in Humanities (used as the reference field of study in estimation)\(^{20}\).

4.2 Expectations formation

I assume that students form their expectations about their admission chances on the grounds of the programs’ past-year admission cutoffs and their own priority score, rather than based on an explicit model for other students’ behavior. This is a natural approach given the public availability of past-year cutoffs. Specifically, to report the expected-utility-maximizing list, students derive their expectations assuming that admission cutoffs follow, from one year to the next, some relationship of the form:

$$\text{cutoff}_{j,2010} = \alpha + \text{cutoff}_{j,2009} + \eta_j \quad \text{with} \quad \eta_j \sim N(0,\sigma^2). \quad (3)$$

Then, taking parameters $\alpha, \sigma$ as given, student $i$’s expectation about her probability to clear the admission cutoff for program $j$ is:

$$P(\text{priority}_i \geq \text{cutoff}_{j,2010}) = \Phi \left( \frac{1}{\sigma} (\text{priority}_i - \alpha - \text{cutoff}_{j,2009}) \right) \quad \text{(4)}$$

For simplicity, to designate $i$’s expectation about her probability to clear the admission cutoff of a program $j$, I use the term admission chance or admission probability.

Figure A.5 in Online Appendix A gives some intuition about the way $\alpha$ and $\sigma$ shape students’ expectations about their admission chances. $\sigma$ captures the level of uncertainty the student perceives to be facing. As $\sigma$ goes to 0, uncertainty decreases: in the limit, cutoffs realization can be predicted to equate their past-year value plus $\alpha$. $\alpha$ is the systematic change in cutoffs expected by the student. A student using $\alpha = 0$ would expect to have .5 probability to beat the 2010 cutoff of a program with 2009 cutoff equal to her own priority score. A student using $\alpha > 0$ behaves as if she thought cutoffs would, on average, increase, so that, if she wants to have a .5 probability of beating the 2010 cutoff of her listed program, she needs to list a program whose 2009 cutoff is below her own priority score. Conversely, a student using $\alpha < 0$ expects to beat the 2010 cutoff of a program with 2009 cutoff equal to her priority score with probability higher than .5.

---

\(^{18}\)The four STEM fields are: Math/Computer Sci.; Physics/Chemistry/Engineering; Biology/Medicine; Earth Sci.

\(^{19}\)Past-year cutoffs are expressed in percentiles of the distribution of priority scores; non-selective programs have cutoff at the 0th percentile.

\(^{20}\)In other words, student $i$ derives flow utility $u_i(\ell) = \beta_{\text{SES}_i} \times \text{distance}_{i,\ell} + \gamma_{\text{SES}_i} \times \text{past-year cutoff}_\ell + \varepsilon_{i,\ell}$ from a program $\ell$ preparing her to receive a BA in Humanities. If program $\ell$ is local (distance$_{i,\ell} = 0$) and non-selective (past-year cutoff$_\ell = 0$), then $i$’s utility for $\ell$ is $u_i(\ell) = \varepsilon_{i,\ell}$. 

17
Distribution. My goal is to recover the distribution of \((\alpha, \sigma)\) in the population of students. I assume that \((\alpha, \sigma)\) is distributed according to a truncated bivariate normal, and I seek to estimate the parameters of this distribution. Formally, assume that:

\[
\begin{pmatrix}
\alpha \\
\sigma
\end{pmatrix} \sim \text{trN}_{[L,U]}(m,S)
\]

where \(m = \begin{pmatrix} m_\alpha \\ m_\sigma \end{pmatrix}, \ S = \begin{pmatrix} s_\alpha^2 & 0 \\ 0 & s_\sigma^2 \end{pmatrix}, \) are the mean and variance matrix of the underlying normal, and \(L = \begin{pmatrix} L_\alpha \\ L_\sigma \end{pmatrix}, \ U = \begin{pmatrix} U_\alpha \\ U_\sigma \end{pmatrix}, \) are the lower and upper truncation bounds.

Based on Figure 1 in Section 3, I set the bounds to \((L_\alpha, U_\alpha) = (-2, 2)\) and \((L_\sigma, U_\sigma) = (0, 1).^{21}\) The discussion of Figure 1 in Section 3 highlighted that students in the top half of the priority score distribution face an environment with relatively less variance than students in the bottom half. Consistently, it is natural to allow the distribution of \((\alpha, \sigma)\) used by students to form their expectations to differ based on one’s priority score.\(^{22}\) In addition, I also allow \(m\) and \(S\) to depend on one’s position within one’s group through the parametrization:

\[
m_\alpha(\text{priority}_i, \text{dist}_{\text{info}_i}) = m_{\alpha,0} + m_{\alpha,1} \times \text{priority}_i + m_{\alpha,2} \times \text{priority}^2_i
+ m_{\alpha,3} \times \text{dist}_{\text{info}_i} + m_{\alpha,4} \times \text{dist}_{\text{info}_i}^2
\]
\[
s_\alpha^2(\text{priority}_i, \text{dist}_{\text{info}_i}) = s_{\alpha,0} + s_{\alpha,1} \times \text{priority}_i + s_{\alpha,2} \times \text{priority}^2_i
+ s_{\alpha,3} \times \text{dist}_{\text{info}_i} + s_{\alpha,4} \times \text{dist}_{\text{info}_i}^2,
\]

where priority\(_i\) denotes \(i\)'s priority score, and dist_{\text{info}_i} is \(i\)'s distance to the beginning of his own application group. I use a similar parametrization for \(m_\sigma\) and \(s_\sigma^2\). To simplify notation, I also denote by \((m,S)\) the parameters to be estimated \(\{(m_{\alpha,k}, m_{\sigma,k}, s_{\alpha,k}, s_{\sigma,k})\}_{k=0,1,2,3,4}\).

Information updates. In this framework, the effect on admission chances of an information update, such as those provided in the Tunisian mechanism is two-fold. First, it allows students to reset some of their expected admission probabilities to 0 or 1. It resets to 0 the perceived probability of admission to programs that are publicly declared to have more vacancies remaining than there are students with higher priority than \(i\) in her group. Second, because \(m\) and \(S\) depend on students’ distance to the beginning of their group, the revelation of information can change the values of \(\alpha\) and \(\sigma\) students use to form their expectations. Indeed, consider the two marginal students \(i\) and \(j\) on each side of a group limit, say the limit between Groups 1 and 2. These two students have virtually the same priority (that is, priority\(_i = \text{priority}_j\)), but the student marginally at the beginning of Group 2 (say, \(j\)) is at distance 0 from the most recent information revelation (that is, dist_{\text{info}_j} = 0), while the student marginally at the end of Group 1 (say, \(i\)) is separated from the the information most recently available to her by the whole Group 1 (that is, 30% of applicants, or dist_{\text{info}_i} = .3). Specification (5) for the components of \(m\) and \(S\) allows the parameters \((\alpha_i, \sigma_i)\) and \((\alpha_j, \sigma_j)\), respectively used by \(i\) and \(j\) to form their expectations, to be drawn from different distributions.

Continuation value \(V_i(0)\). In the Tunisian sequential design, Group-1 and Group-2 students who fail to be assigned to any of their listed choices are pooled at the top of the next group and allowed to participate again in the application process. At that time, they can only pretend to alternatives still available in the

---

\(^{21}\)These bounds are chosen with two concerns in mind. First, they should be wide enough not to constrain the estimates. Second, as will be made clear in the Section 5.3, tighter bounds allow for more precise estimation at a lower simulation cost.

\(^{22}\)Recall that Figure A.1 in Online Appendix A shows that, conditional on priority score, high- and low-SES students apply to programs with similar selectivity levels.
choice set of this next group. This “second chance” affects students’ option value of being rejected from all their listed choices—in Equation (1), $V_i(0)$. I assume that a student $i$ in Group 1 or 2 computes $V_i(0)$ as follows. Let $\text{groupCutoff}_i$ denote the position in the priority ranking of the division between $i$’s application group and the next. Student $i$ can form expectations about the choice set she would face at the time of her second chance if she were to use it by using $\text{groupCutoff}_i$ instead of $\text{priority}_i$ in Equation (4). Then, $V_i(0)$ is the expected value of the program with highest expected utility, based on these second-chance expectations. $i$’s expected probability to use her second chance ($\pi_i$ in Equation (1)) is determined by her expectations about her admission probabilities to her listed choices.

5 Identification and estimation

The identification strategy takes full advantage of local incentives for truth-telling induced by a sequential implementation of the DA. It proceeds in two steps. First, I argue that the design of the Tunisian mechanism creates incentives for a subset of students to simply list their most-preferred programs. For these students, preferences can be recovered within a standard discrete choice framework. I argue that the utility parameters recovered from this subset of truthful students actually represent the preferences of the full sample. Part of this extrapolation argument relies on assumptions about the utility function, but I show that the structure of the sequential application procedure and of the subset of truthful students makes the extrapolation very reasonable. Second, expectations parameters are identified as those rationalizing other students’ observed application lists, given the identified preferences.

5.1 Preferences

Truthful students. Given the information revelation made before each group submits applications, the first-ranked student in each group is faced with making a choice under perfect information. If she knows she is ranked first in the group, she knows she has probability 1 to be assigned to the first-ranked element of her list (as long as she lists a program that has not been publicly declared full). It is therefore strictly dominant for her to list her most-preferred program first in her application ranking. And it is weakly dominant for her to also list her second to tenth most-preferred programs. Because students may apply to up to ten programs and because most programs have more than one vacancy, not only the first-ranked student, but a subset of applicants at the top of each group have incentives to truthfully report their most-preferred programs. Proposition 2 and Condition 1 stated in Section 2 give a sufficient condition for students to truthfully report their most-preferred programs. Proposition 2 and Condition 1 stated in Section 2 give a sufficient condition for students to truthfully report their most-preferred programs. They imply that, when going down along the priority ranking within a group, students will truthfully report their most-preferred programs as long as they think they have probability 1 to have a seat in one of their ten most-preferred programs. Later in this section, I provide empirical evidence that students at the top of each group indeed list their most-preferred programs among those publicly declared to still have seats.

Extrapolation. If truthful students were a random subset of the applicant sample, parameters recovered from the former would straightforwardly represent the preferences of the latter. Recall that I call “truthful” students who are given the incentives to reveal their preferences by their position at the top of their application group. While these students are locally representative of students around group cutoffs (see Section 3), they are a priori not representative of the whole sample. However, Table A.1 in Online Appendix A shows that the subset of truthful students is representative of the full applicant sample in terms of gender, socioeconomic status, and home location, in the sense that these characteristics have similar support and variation in both sets of students. This is, by construction, not the case for high-school performance, since truthful students are sampled from three points of the priority distribution. I rely on the assumed continuity of the utility function to extrapolate the heterogeneity in preferences across high-school performance from truthful students to the full sample. Two main elements makes such extrapolation credible in the Tunisian

---

23The possibility to be tied in the priority order may encourage students to list choices beyond the very first rank.
setting. First, the sequential design yields truthful students at three very different high-school performance levels, rather than a single one if the assignment procedure was implemented in its standard one-phase form. This means that, rather than assuming that preferences for college programs are independent of high-school performance, as would be necessary if only students at the very top of the one-group cohort had the incentive to be truthful, preferences of students with medium performance are predicted using the preferences of truthful students with both higher and lower academic performance than them. Second, the data, which contain students’ scores at the national exam in eight different subjects, allow for a flexible parametrization of the effect of academic performance on preferences for programs. Hence, extrapolation to the full applicant sample is not simply done linearly from three points in the composite priority score, but accounting for heterogeneity in the multiple dimensions that enter this priority score. Specifically, I let students’ preferences for the different fields of study (e.g. math) depend separately on their own high-school performance in the field (i.e. math), as well as in other closely-related fields (i.e. other STEM fields), and in more distant fields (e.g. non-STEM fields), to account for individual comparative advantages in studying one subject vs another. I also let students’ preferences for the type of degree pursued depend on their end-of-high-school performance. To test the extrapolation argument, I estimate utility parameters using truthful students are the top of Groups 1 and 3 only, and then predict choices for truthful students at the top of Group 2. Predictions and the data are compared in Section 6.1.

**Estimation.** I estimate utility parameters by maximum likelihood, using students at the “top” of each group as my estimation sample. Details about the likelihood functions are shown in Online Appendix D. The choice of the “top” is viewed as a bandwidth choice problem, solving a bias-variance tradeoff. The preferred estimates, which are shown in Tables 5 and 6 in Section 6.1, are based on an estimation sample using all Group-1 and Group-3 students with within-group rank below 200 (that is, the top 5% of Groups 1 and 3).

**Empirical evidence of “top” students being truthful.** In Online Appendix A, I present three pieces of empirical evidence suggesting that students at the top of application groups, which I use as my estimation sample for utility parameters, truthfully report their most-preferred programs —that is, that my estimates are unlikely to be biased by the presence of students misreporting their preferences. I briefly summarize this evidence here. First, the regression-discontinuity analysis in Section 3 shows that students understand and use the information they are provided at the beginning of each group. The second piece of evidence relies on another unique feature of the Tunisian procedure: a reassignment round. After students of all three groups have been assigned but before the new academic year starts, students may, without foregoing their initial assignment, submit a new ordered list of up to four programs that they would prefer to attend over their assigned match. No precise procedure is explicitly defined regarding the processing and approval of requests. It is generally understood that relative priority among students is preserved in the reassignment round and that approval depends on the ability of the requested program to welcome an additional student. Online Appendix Table A.2 shows that ‘Top’ students apply for reassignment at a significantly lower rate than other students (17.6 vs 24.6%) and when they do, they submit fewer requests (2.7 vs 3 programs included in the reassignment list). Most importantly, 84% of the programs top students request are outside their initial choice set, that is, had already been declared full when the students applied in the main procedure. On the contrary, 54% of requests submitted by other students are within their initial choice set. Finally, Online Appendix Figure A.4 shows that, within group, the frequency at which students apply to the programs listed by the very first few students in their group remains high throughout the first 200-250 students in the group, and only starts decreasing after then. This suggests that applicants ‘omitting’ the highest-utility programs from their list can only be found past the 200-250 students in a group.

---

24To test the extrapolation argument, I estimate utility parameters using truthful students are the top of Groups 1 and 3 only, and then predict choices for truthful students at the top of Group 2. As a robustness check, I also show estimates using different bandwidth sizes in Online Appendix A. Online Appendix tables also illustrate the robustness of results to using only students’ first few listed choices (e.g. the first 5 choices) rather than all of them.
5.2 Can truth-telling explain all application portfolios?

Given preferences parameter estimates, one can simulate the application lists that students would choose if they all listed their most-preferred programs among those that have not been publicly declared full. Comparing these simulated lists to those observed in the data provides a simple intuition of the strategy that I will use to recover expectations-formation parameters.

Figure 3 is a bin-plot of the selectivity level (in terms of past-year cutoff) of students’ first- and sixth-listed choices as a function of their priority score. Each graph shows both the data (in black), and simulations based on the estimated utility parameters (in red) under the assumption that every student is truthful. The thickest dotted line with cross-shaped markers show the median selectivity levels of students’ first (resp. sixth) choice along the priority ranking. The thinner dotted lines show other percentiles of this distribution.

While simulations under the truth-telling assumption fit the data at the very beginning of each group, the discrepancy between application profiles in the data and under the truth-telling assumption becomes large as priority decreases within each group. Figure 3 shows three patterns that will be used to pin down the parameters of the expectations-formation model:

1. **Range** — In the data, there is a large range in the selectivity levels of students’ listed choice within groups. This range is larger at priority scores at the end of each group, than at priority scores at the beginning of the group. This range is also larger for choices ranked further down in the application list than for choices ranked earlier in the application list, particularly for priority scores at early in a group. Simulations under the assumption that students truthfully list their most preferred choices do not fully reproduce the variance observed in the data, particularly for Groups 1 and 2.

2. **Slope along priority ranking** — In the data, (each percentile of) the distribution of selectivity levels slopes down as a function of priority ranking. Simulations under the assumption that students report truthfully their ten most-preferred program generate much flatter patterns than what is observed in the data, particularly for Groups 1 and 2.

3. **Slope down application lists** — In the data, the range of selectivity levels move slightly down along the selectivity scale as we look at choices ranked further down in students’ application list. This slope of selectivity levels down application lists is notably not as steep as the slope of selectivity levels down the priority ranking (Pattern 2).

5.3 Expectations

I estimate the model of expectations formation using a two-step method-of-moments approach, in the spirit of the estimators proposed by Ackerberg (2009) and Bajari, Fox, and Ryan (2011), and used by Nevo, Turner,
and Williams (2016). First, I solve the application problem, given preferences, for an array of possible pairs \((\alpha, \sigma)\), which I call types here. Then, I recover the parameters governing the distribution of types in the population by matching a set of data moments to the equivalent moments predicted by the model. Like in Bajari et al. (2011) and Nevo et al. (2016), I implement the two-step method by using the assumed distribution of unobserved types to write the empirical moments as a mixture of conditional moments. An important distinction from these papers is that I want to allow the distribution of unobserved heterogeneity to depend on a set of (continuous) observed individual characteristics — priority and position within one’s group. I use insights from numerical integration to devise an estimation method in which the model only needs to be solved a finite number of times, outside the optimization routine.

**Moments.** I estimate parameters by minimizing the distance between moments of the data and moments predicted by the model. Formally:

\[
(m, \hat{S}) = \arg \min_{(m, S)} \ m(m, S)W^{-1}m(m, S)
\]

where \(m(m, S) = \hat{m}_{\text{data}} - m_{\text{model}}(m, S)\), and \(W\) is a weighting matrix.\(^{25}\)

The choice of moments is based on the discussion of Section 5.2: the parameters of the expectations-formation model should be able to explain the patterns of the application distribution that cannot be generated by truthful behavior. Formally, I use two sets of moments:

1. The mass of individuals in decile \(d\) of the priority distribution whose \(ch^{th}\) listed choice has previous-year selectivity below level \(\kappa\):

\[
m_{1,d,ch,\kappa}^{\text{data}} = \frac{1}{N_d} \cdot \sum_{i \in d} 1 \{\text{cutoff}_{i,(ch),2009} \leq \kappa\}
\]

where \(\{i \in d\}\) denotes the set of students in decile \(d\) of the priority distribution, \(N_d\) is the number of such students, and \(\text{cutoff}_{i,2009}\) is the 2009 (previous year) cutoff for program \(j\).\(^{26}\)

2. The mass of individuals in decile \(d\) of the priority distribution for whom the difference in previous-year selectivity between their first- and \(ch^{th}\)-listed choices is below level \(\chi\):\(^{27}\)

\[
m_{2,d,ch,\chi}^{\text{data}} = \frac{1}{N_d} \cdot \sum_{i \in d} 1 \{(\text{cutoff}_{i,(ch),2009} - \text{cutoff}_{i,(1),2009}) \leq \chi\}
\]

The first set of moments summarizes the across-individual range of selectivity levels students list in their application; while the second set of moments summarizes the within-individual range of selectivity levels each student lists in his/her application.

**Computation.** Denote the PDF of the truncated normal distribution by:

\[
(\alpha, \sigma) \mapsto \phi_{[L, U]}(\alpha, \sigma; m, S) = \frac{\phi(\alpha, \sigma; m, S)}{Pr_{m,S}(L \leq X \leq U)}
\]

where \(\phi(\cdot, m, S)\) is the PDF of the (not truncated) bivariate normal distribution with mean vector \(m\) and variance-covariance matrix \(S\); \(Pr_{m,S}(L \leq X \leq U)\) is the probability for a variable \(X \sim N(m, S)\) to take a value between \(L\) and \(U\).

The model-predicted moments analogous to \(m_{1,d,ch,\kappa}^{\text{data}}\) defined above can be written as a linear function of

\(^{25}\)In the estimation, I set \(W\) equal to the identity matrix.

\(^{26}\)For each decile down the priority ranking, I use \(\kappa\) corresponding to the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of selectivity levels (as illustrated in Figure 3, for instance). For \(d = 1, \ldots, 10\), \(ch = 1, 3, 6, 9\), and \(\kappa\) taking five values for each group, this yields \(10 \times 4 \times 5 = 200\) moments.

\(^{27}\)Again, for each decile down the priority ranking, I use \(\chi\) corresponding to the 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of changes in selectivity level (so that \(\chi\) is a difference of percentiles). For \(d = 1, \ldots, 10\), \(ch = 3, 6, 9\), and \(\chi\) taking five values, this yields \(10 \times 3 \times 5 = 150\) moments.
the density of \((\alpha, \sigma)\):

\[
m_{1, d, ch, \kappa}^{\text{model}}(m, S) = \mathbb{P}^{\text{model}}\left[\text{cutoff}_{t_i(ch), 2009} \leq \kappa \mid i \in d\right]
\]

\[
= \frac{1}{N_d} \sum_{i \in d} \left\{ \int_{[L, U]} \mathbb{P}^{\text{model}}\left[\text{cutoff}_{t_i(ch), 2009} \leq \kappa \mid (\alpha, \sigma) \right] \times \phi_{[L, U]}(\alpha, \sigma; m, S) \, d(\alpha, \sigma) \right\}
\]

\[
= \frac{1}{N_d} \sum_{i \in d} \Pr_{m(i), S(i)}(L \leq X \leq U) \left\{ \int_{[L, U]} \mathbb{P}^{\text{model}}\left[1_{1, i, ch, \kappa}(\alpha, \sigma) \times \phi(\alpha, \sigma; m(i), S(i) \mid (\alpha, \sigma) \right] \right\}
\]

where \(\Pr_{m(i), S(i)}^{\text{model}}(\alpha, \sigma)\) is the value of the model-predicted probability that \(i\)'s \(ch^{th}\) listed choice has past-year cutoff lower than \(\kappa\) conditional on \(i\) using the pair \((\alpha, \sigma)\) to form her expectations, and where I use the notation \((m(i), S(i))\) in the summand to emphasize the fact that the mean and the variance-covariance matrix of the normal distribution depend on \(i\)'s priority score. An analogous equation holds for the second set of moments \(m_{2, d, ch, \chi}^{\text{model}}\).

At this stage, two points need to be highlighted. First, the parameters to be estimated, \((m, S)\), only enter the normal weights \(\phi(\alpha, \sigma; m(i), S(i))\) inside the integral and the scaling factor \(\Pr_{m(i), S(i)}(L \leq X \leq U)\) outside the integral. Both of these terms are quick and easy to compute for a bivariate normal. On the contrary, given a pair \((\alpha, \sigma)\), the computationally more demanding terms \(\Pr_{m(i), S(i)}^{\text{model}}(\alpha, \sigma)\) inside the integral do not depend on \((m, S)\). Second, the integral \(I_{1, i, ch, \kappa}(m, S)\) can be computed numerically using quadrature methods if the mapping \((\alpha, \sigma) \mapsto \Pr_{m(i), S(i)}^{\text{model}}(\alpha, \sigma)\) is known at the nodes of the quadrature. Importantly, the nodes used in standard quadrature methods depend only on the bounds \((L, U)\) of the domain of integration, but not on the parameters \((m, S)\).

This means that I can solve the estimation problem (6) in two steps. First, I compute predicted probabilities \(\Pr_{m(i), S(i)}^{\text{model}}(\alpha, \sigma)\) and \(\Pr_{2, i, ch, \chi}^{\text{model}}(\alpha, \sigma)\) for an array of nodes \((\alpha, \sigma)\) for all individuals \(i\) and all \((ch, \kappa, \chi)\). Second, I numerically approximate the objective function in the estimation problem (6) and solve for \((m, S)\).

In the first step, I use quadrature rules and the method proposed by Smolyak (1963) and Heiss and Winschel (2008)\(^{28}\) to determine a sparse two-dimensional grid \(\{x_t = (\alpha_t, \sigma_t)\}_{t=1}^{T}\). For each point \(x_t\), I estimate the conditional probabilities \(\Pr_{m(i), S(i)}^{\text{model}}(x_t)\) and \(\Pr_{2, i, ch, \chi}^{\text{model}}(x_t)\) for all \(i\)'s and \((ch, \kappa, \chi)\)'s by simulation. In each simulation, I construct application lists using Chade and Smith (2006)'s Marginal Improvement Algorithm.\(^{29}\)

In the second step, I solve (6) by approximating \(I_{1, i, ch, \kappa}(m, S)\) (and analogously for \(m_{2, d, ch, \chi}^{\text{model}}\)) by:

\[
I_{1, i, ch, \kappa}(m, S) \approx \sum_{t=1}^{T} \Pr_{m(i), S(i)}^{\text{model}}(x_t) \times \phi(x_t; m, S) \times \omega_t
\]

where \(\{\omega_t\}_t\) is the set of weights associated with the quadrature nodes. Online Appendix D gives more details about the estimation method.

Inference. To calculate standard errors, I use a block-resampling procedure (Lahiri, 2003). I sample the data by student with replacement, keeping the whole application list for each of the students drawn. I use 100 bootstrap samples and for each sample, I proceed in two steps. First, I estimate utility parameters using students in the bootstrap sample with priority falling within the estimation bandwidth discussed in

---

\(^{28}\)While I use the method within a two-step method-of-moments estimation strategy, Heiss and Winschel (2008) discuss the implementation and the performance of this method for approximating likelihood functions, and provide the values of nodes \(\{x_t\}_t\) and weights \(\{\omega_t\}_t\) of the quadrature.

\(^{29}\)When a student’s probabilities to clear cutoffs are, conditional on her priority score, independent across programs, the Marginal Improvement Algorithm finds the solution to Problem (1) (Chade & Smith, 2006). This is the case in this setting, given the distributional assumption in (3), and the fact that students know their priority score at the time of applying. However the Marginal Improvement Algorithm is not guaranteed to find the exact solution to Problem (1) when these probabilities are not independent, for example because students do not know their priority score at the time of application as in the setting of Ajayi and Sidiabé (2017).
Section 5.1. Then, I use the bootstrap estimates to re-compute the predicted moments, and I re-estimate the expectations-formation parameters by matching predicted moments to bootstrap-sample moments from the data. Next, to determine confidence intervals in Section 7, I repeat the counterfactuals analysis for each bootstrap sample.

Identification. Because the parameters to be recovered enter the decision problem in a nonlinear way, it is hard to pinpoint which variation in the data identifies each of the expectations-formation parameters. However, an illustration of how different values of \((\alpha, \sigma)\) generate very different distributions of application profiles in the population helps understanding how the parameters in \(m\) and \(S\) are pinned down.

As an example, the graphs in Figure 4 shows moments of the predicted distribution of selectivity levels of students’ third listed choice as a function of priority ranking, conditional on given values of \((\alpha, \sigma)\). The top panel shows the median selectivity levels of students’ third listed choice as a function of their priority ranking; the middle and bottom panels show the 80th and 20th percentiles of this distribution, respectively. Panels on the right-hand side (RHS) show listed selectivity levels when students use a relatively large value of \(\sigma\) to form their expectations \((\sigma=.44)\), while panels on the left-hand side (LHS) assume students use a relatively small value of \(\sigma\) to form their expectations \((\sigma=.05)\). Given this value of \(\sigma\), each subfigure show listed selectivity levels for a range of values of \(\alpha\) that students could be using to form their expectations—from a large and negative value \((\alpha=-.65)\) to a large and positive value \((\alpha=1.16)\). Figure 5 follows the same logic, but illustrates how \(\alpha\) and \(\sigma\) affect the second set of moments used in the estimation. It represents the distribution of slope in selectivity within individual lists. Each plots, as a function of priority score, the moments of the selectivity distribution generated when students use a small value of \(\sigma\) to form their expectations; the bottom panel when they use a large value of \(\sigma\). Each panel displays distributions for various values of \(\alpha\).

The top panel of Figure 4 shows that, regardless of the distribution of \(\sigma\) the priority ranking \((\sigma\) close to 0 as in the RHS, or relatively large, as in the LHS), large and positive values of \(\alpha\) are needed to generate the level of selectivity observed in the data for choices made by students at the end of Group 1, while relatively large negative values of \(\alpha\) are required to generate the level of selectivity observed in the data for choices made by students in Group 3. Given the intuition given about \(\alpha > 0\) and \(\alpha < 0\) in Section 4.2 and Online Appendix Figure A.5, this is consistent with the observation of Figure 3, which shows that in the data, the application lists of Group-3 students are close to what truthful students would submit, while application lists of end-of-Group-1 students are much less selective than what would be induced by truthful behavior. This identifies a strongly negative slope \(m_{\alpha,1}\), as well as a relatively large and negative baseline \(\alpha\) at the bottom of the priority ranking \((m_{\alpha,0})\).

The top panel of Figure 4 also shows there are two ways to generate the observed slope in selectivity of choices within groups: (i) the mean \(\alpha\) decreasing along the priority ranking (as in the top right panel where the empirical median intersects successively the medians simulated from decreasing \(\alpha\)); or (ii) individuals expecting cutoffs to change very little \((\sigma\) close to 0, as in the top left panel).

Comparing the top to the middle and bottom panels of Figure 4 shows that each moment of the distribution of listed selectivity levels slopes down along the priority ranking and within group. Note, however, that the set of downward slopes in \(\alpha\) (both along the priority ranking and within groups) holding \(\sigma\) constant, and of values of \(\sigma\) given \(\alpha\) that can generate the slope down in the moment of the distribution of listed selectivity is not the same for all moments. As shown by Figure A.6 in Online Appendix A, these sets also do not perfectly coincide for different listed choices, highlighting that the parameters of interest affect the moments of the selectivity distribution in a way that is not collinear across choices.

Figure 5 shows that the range of heterogeneity across individuals’ within-list change in selectivity is also dramatically affected by the values of \(\alpha\) and \(\sigma\) that students are using. Figure 5 highlights that matching the observed range of heterogeneity in within-list selectivity change requires heterogeneity in \(\alpha\) and \(\sigma\) not only across but also within priority levels, especially in the bottom half of the priority ranking.

---

30 This first step gives the standard errors and confidence intervals shown in Tables 5 and 6, and Figure 6 in Section 6.1.
Figure 4: Distribution of the selectivity of choices in the data and conditional on \((\alpha, \sigma)\)

(a) Median

(b) 80th percentile

(c) 20th percentile

Note: Conditional on a pair \((\alpha, \sigma)\), the predicted distribution of the selectivity of students' listed choices is obtained by simulation using the utility parameters shown in Tables 5 and 6. Shown percentiles are taken within 500-student bin.
Figure 5: Range of within-list change in selectivity across choices in the data and conditional on $(\alpha, \sigma)$

20th and 80th percentiles of relative selectivity of choice 9; $\sigma=0.056$

Note: Conditional on a pair $(\alpha, \sigma)$, the predicted distribution of selectivity differences is obtained by simulations using the utility parameters shown in Tables 5 and 6. Percentiles are taken within 500-student bin.

While each parameter is not identified by an exclusive subset of the moments, Figures 4 and 5 show that the estimator pins down the relative contribution of heterogeneity, both within and across priority levels, in perceived trend and perceived uncertainty (that is, the parameters in $S$ and $m$) by matching moments that are affected by the parameters of interest in a non-collinear way.
6 Parameters estimates

6.1 Preferences estimates

In Table 5, the statistically significant negative coefficient on Distance means that, on average, students prefer attending a program closer to home than further. The estimate coefficients associated with gender and SES interacted with an indicator for whether the program is in the student’s home region suggest that there is no systematic heterogeneity in disutility for traveling across demographics. However, the estimated variances for the random coefficients on distance and the home dummy suggest that there is some idiosyncratic heterogeneity in students’ disutility from traveling.

A positive coefficient on Past-year cutoff means that students value program quality, as measured by the priority ranking of the marginally-admitted student in the previous year. An increase in Past-year cutoff by .1 means that the 2009 marginal admit’s priority score is higher by 10 percentile points. A positive coefficient on squared Past-year cutoff means that the marginal value of an increase in program selectivity increases with the program level of selectivity. In other words, students’ willingness to travel to attend a program with marginally higher quality increases as quality gets higher.

In Table 6, a positive coefficient on a degree indicator means that, on average, students prefer the degree (BS or advanced degree) to the reference degree (BA). A higher high-school performance correlates with stronger preferences for an advanced degree. A positive coefficient on a field indicator means that, on average, students prefer the field to the reference field (Humanities). All field-dummy coefficients being positive except the Law one means that Law is the least-preferred field for average-performing male students, while Humanities is the second least-preferred. Comparison of field-dummy coefficients show that STEM fields are preferred over non-STEM fields, and that Math/Computer Science is the most-preferred field of study. This is not surprising given that students in the sample graduated from high-school with a Math major. None of the coefficients that allow for heterogeneous preferences across gender are statistically significant, although their signs seem to suggest that, all other things equal, female students have a weaker preference for Math/Comp.Sci. and Physics/Chemistry/Engineering, and a stronger preference for Health and Life Sciences and Earth Sciences relative to males. For students of both sexes, Earth Science is the least-preferred STEM field. High-school performance appears to be a strong driver of preferences for the different fields of study. A higher performance in non-STEM fields correlates with a weaker preference for STEM fields and more mathematically-oriented non-STEM fields, such as Economics. On the contrary, a higher performance in STEM fields correlates with a stronger preference for STEM fields, except for the least-preferred STEM field, Earth Science.

Model fit. Figure 6 illustrates the good fit of the estimated model for students at the top of each group. The black lines are plots from the data, the colored lines show predictions based on estimates from Tables 5 and 6. The thick solid lines in each panel show the median selectivity of listed choices 1 to 10 observed in the data for top students in each group. The gray bands show 95%-confidence intervals. To illustrate the fit beyond the median, the thinner dotted lines the 30th and 70th percentiles of the distribution of listed choices within the top or each group. Except for choices 8 to 10 of students at the top of Group 1, the estimated model manages to reproduce the selectivity distribution of top students’ applications, including for students at the top of Group 2, who were not used in the estimation of utility parameters.

---

31 High-school performance in STEM and non-STEM are standardized to have mean 0 and standard-deviation 1 in the population of students graduating from high-school with a Math major.

32 This part of the data not being perfectly matched is not a concern, as explained by Figure 6 and Figure A.2 in Online Appendix A. Figure 6 suggests that in the data, the programs listed 8 to 10 by students at the top of Group 1 have much lower past-year cutoffs than those ranked 1 to 7 by these students. This drop in selectivity is not predicted by simulations based on estimated utility parameters. However, Figure A.2 shows that very few students at the top of Group 1 actually rank more than 7 choices. Figure 6 shows the distribution of selectivity of the kth listed choice, conditional on students ranking at least k choices. Beyond choice 7, bin-averages for Group 1 shown in Figure 6 are therefore based on few observations.
Table 5: Estimated preferences parameters

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (100km)</td>
<td>-1.62</td>
<td>0.23</td>
</tr>
<tr>
<td>Squared</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.83</td>
<td>0.13</td>
</tr>
<tr>
<td>Home</td>
<td>-0.4</td>
<td>0.33</td>
</tr>
<tr>
<td>$\times$ high SES</td>
<td>0</td>
<td>(0.3)</td>
</tr>
<tr>
<td>$\times$ female</td>
<td>0.04</td>
<td>0.29</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>2.06</td>
<td>(0.3)</td>
</tr>
<tr>
<td>Past-year cutoff</td>
<td>1.15</td>
<td>0.46</td>
</tr>
<tr>
<td>$\times$ high SES</td>
<td>0.77</td>
<td>(0.49)</td>
</tr>
<tr>
<td>Squared</td>
<td>3.89</td>
<td>(0.48)</td>
</tr>
<tr>
<td>Past-year cutoff $\times$ Dist (100km)</td>
<td>0.71</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Degree: 3-yr, fundamental track</td>
<td>0.38</td>
<td>(0.14)</td>
</tr>
<tr>
<td>High-school perf.</td>
<td>0.22</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Degree: advanced</td>
<td>2.67</td>
<td>(0.21)</td>
</tr>
<tr>
<td>High-school perf.</td>
<td>0.77</td>
<td>(0.22)</td>
</tr>
<tr>
<td>PseudoObs.</td>
<td>3,529</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>415</td>
<td></td>
</tr>
</tbody>
</table>

Bootstrap std. err. in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Estimated preferences parameters (ctnd.)

<table>
<thead>
<tr>
<th></th>
<th>$\times$ STEM score</th>
<th>$\times$ non-STEM score</th>
<th>$\times$ female</th>
<th>$\times$ field score</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>1.42**</td>
<td>1.09**</td>
<td>-0.72*</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.9)</td>
<td>(0.61)</td>
<td>(0.46)</td>
<td>(0.89)</td>
<td></td>
</tr>
<tr>
<td>Tertiary</td>
<td>3.49</td>
<td>0.82*</td>
<td>0.1</td>
<td>-1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.55)</td>
<td>(2.97)</td>
<td>(1.43)</td>
<td>(8.22)</td>
<td></td>
</tr>
<tr>
<td>Soc.Sci</td>
<td>0.63</td>
<td>1.59</td>
<td>0.14</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(27.79)</td>
<td>(22.8)</td>
<td>(3.54)</td>
<td>(10.35)</td>
<td></td>
</tr>
<tr>
<td>Econ/Mgmt</td>
<td>2.73****</td>
<td>0.36</td>
<td>-0.18</td>
<td>-0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(0.58)</td>
<td>(0.41)</td>
<td>(0.84)</td>
<td></td>
</tr>
<tr>
<td>Law</td>
<td>-37.06</td>
<td>1.08*</td>
<td>-2.69**</td>
<td>38.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(21.23)</td>
<td>(3.16)</td>
<td>(9.35)</td>
<td>(9.93)</td>
<td></td>
</tr>
<tr>
<td>Physics/Engin.</td>
<td>3.34***</td>
<td>0.96</td>
<td>-0.37</td>
<td>-0.65</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.58)</td>
<td>(0.4)</td>
<td>(0.84)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Health &amp; Life Sci.</td>
<td>2.45****</td>
<td>0.28</td>
<td>-0.86***</td>
<td>0.59</td>
<td>0.26*</td>
</tr>
<tr>
<td></td>
<td>(0.9)</td>
<td>(0.64)</td>
<td>(0.44)</td>
<td>(0.85)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Earth Sci.</td>
<td>1.39***</td>
<td>-0.78**</td>
<td>-0.96**</td>
<td>0.36</td>
<td>0.36***</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.51)</td>
<td>(0.47)</td>
<td>(0.89)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Math</td>
<td>3.48***</td>
<td>0.14</td>
<td>-0.81</td>
<td>-0.93</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.92)</td>
<td>(0.64)</td>
<td>(0.43)</td>
<td>(0.86)</td>
<td>(0.17)</td>
</tr>
</tbody>
</table>

Bootstrap std. err. in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
6.2 Expectations parameters

Estimated values for $m$ and $S$ are illustrated by Figure 7. They point to three main facts. First, the top LHS panel shows there is a strongly negative slope in $m_\alpha$ along the priority ranking, which was expected from the observation of Figures 3 and 4. Figure 3 showed that the observed application lists of Group 3 students are quite close to profiles that would be chosen by students simply reporting their most preferred programs. Figure 4 established that the observed application lists for Group 3 students, close to truthful application lists, are consistent with students behaving as if they expected that cutoffs would, on average, decrease. On the contrary, Figure 3 revealed that the observed application lists of Group 1 students, especially in the second half of the group, have much lower selectivity levels than what we would observe for truthful students. Figure 4 confirmed that the observed application lists for Group 1 students can only be the result of them behaving as if they expected that cutoffs would, on average, increase. Second, the top RHS panel shows a strong increase in perceived uncertainty $m_\sigma$ as we go down the priority ranking. The steepest increase happens between the bottom of Group 1 and the top of Group 2, while the increase is smoother through the second part of the priority ranking. This is interesting because consistent with the fact that, as established by Figure 1, students with lower priority face an environment with much more uncertainty than students at the top of the priority ranking. The LHS panel of Figure 1 actually suggests that the year-to-year variance in cutoffs sharply increase as we reach the top third of the score distribution. Third, the bottom two panels show there is a strong increase in heterogeneity in both perceived trend and perceived uncertainty as priority decreases. The level of heterogeneity in perceived trend and uncertainty in relatively low among Group-1 students, meaning that they all more or less perceive the environment in the same way. On the contrary, larger heterogeneity in Groups 2 and 3 means that they behave as if their individual perceptions of the environment differed substantially from each other’s.
Figure 7: Estimated $m$ and $S$ as a function of priority

Model fit. The correlation between empirical moments and their fitted counterparts is equal to .97. Figure 8 further illustrates the fit of the estimated model. It shows that the model is able to predict well the empirical distribution of the selectivity of listed elements, even for those not directly used in the estimation (e.g., choices 2 and 7).

Figure 8: Predicted vs. data distribution of selectivity of choices (2010)
7 The Value of Information and Fallback Options

In this section, I evaluate the effects of sequencing the implementation of the restricted-list DA mechanism on students’ applications, assignments, and the utility they receive, \textit{ex post}, from their assignment.

I do so by simulating students’ applications\footnote{In each simulation, I construct students’ application lists using Chade and Smith (2006)’s Marginal Improvement Algorithm.} and assignments under four multiple-phase scenarios — dividing the cohort in two, three, four, or five equally-sized groups, by order of priority — and comparing them to counterfactual applications and assignments obtained under two benchmarks. My first benchmark is the standard one-phase implementation of the restricted-list DA: all students submit applications simultaneously, and are administratively assigned to leftover seats if they get rejected from all their listed choices. As my second benchmark, I use the perfect-information setting in which information about vacancies is publicly updated after every single assignment. This corresponds to a limit \( N \)-phase scenario, where \( N \) is the total number of students in the population; given my assumptions, it yields the same final match as the unrestricted-list DA. The priority ranking is kept fixed across scenarios; and parameters governing preferences and expectations formation (that is, \( \beta, \gamma, \nu, \) and \( m_\alpha, s_\alpha, m_\sigma, s_\sigma \)) are treated as being policy-invariant. To compare scenarios, I use the uniformly weighted sum of \textit{ex post} utilities as my measure of student welfare:

\[
W = \frac{1}{N} \sum_{i=1}^{N} E\left[u_{i,\mu(i)} + \varepsilon_{i,\mu(i)}\right]
\]

where \( \mu(i) \) denotes student \( i \)'s assignment, and the expected value \( E \) is estimated by simulations over utility parameters \( \varepsilon \) and \( \lambda \), and expectations-formation parameters \( \alpha \) and \( \sigma \). Let \( \Delta W_{PI,1} \) denote the difference in expected student welfare between the perfect-information situation and the single-phase scenario. \( \Delta W_{PI,1} \) represents the potential welfare gains that can be achieved by using a sequential implementation of the DA instead of a single-phase procedure.

7.1 Expected Gains From Sequential Implementation

Counterfactual simulations show that, under perfect information, expected student welfare is higher by the equivalent of reducing the average student’s distance to her benchmark single-phase assignment by 70.3km (43.6 miles; 95%-confidence interval: [57.6; 133.2]) — I denote: \( \Delta W_{PI,1} = 70.3 \text{km} \). As a reference, the median distance traveled in the data is 100km; and the average utility decrease between a student’s most-preferred program and her second most-preferred program is equivalent to an 58km-increase in the student’s distance to her most-preferred program (see Figure A.7 in Online Appendix A). This difference increases to 93, 115, 132, and 144 km if we consider the average utility decrease between a student’s most-preferred program and her third, fourth, fifth, and sixth most-preferred programs, respectively.

Figure 9 shows that a small number of phases is enough to achieve a very large share of the potential gains \( \Delta W_{PI,1} \). Using a two-phase procedure, instead of the standard single-phase implementation, allows to realize 68% of the welfare gains that would be achieved by moving from one phase to perfect information. This number increases to 89% and 98% when using three or five phases, respectively.

Figure 10 shows expected welfare gains as a function of students’ position in the priority ranking. The solid red plot shows difference in expected student welfare between the perfect-information setting and the single-phase benchmark, while the black and gray plots show welfare gains for an implementation in two, three, or five phases, respectively. The figure shows four main interesting facts. First, the red plot shows that, relative to a perfect-information setting, using a single-phase restricted-list DA in an environment where students face uncertainty about their admission chances creates utility losses for students at all priority levels, including early on in the priority ranking. The gray dashed plot being almost confounded with the perfect-information plot shows that a five-phase implementation restores virtually all of these post-utility losses, which is consistent with what Figure 9 suggested. Second, the two-phase implementation restores a very large share of post-utility losses in the second half of the priority ranking but a much smaller share of the...
Figure 9: Expected change in average assignment utility, relative to one-phase benchmark

Share of potential welfare gains \( \Delta W_{PI,1} \) achieved as a function of the number of phases implemented. Recall that \( \Delta W_{PI,1} \) denote the difference in expected student welfare between the perfect-information situation and the single-phase scenario.

Read: Using a two-phase procedure, instead of the standard single-phase implementation, allows to realize 68% of the welfare gains that would be achieved by moving from one phase to perfect information.

losses experienced by the first half of the priority ranking. The marginal gains of a third phase virtually all accrue to students in the first half of the priority ranking. Third, while the two-phase implementation features a revelation of information about available seats to students in the bottom half of the priority ranking, the gains from a two-phase implementation, relative to a standard single-phase procedure, increase steadily throughout the top half of the priority ranking to peak at the beginning of the second half. Finally, students in the third quarter of the priority ranking get higher utility from their assignment under the two-phase implementation than from their perfect-information assignment. A three-phase implementation bring these students’ post utility down to its perfect-information level, but instead gives an edge to students at the beginning of the second third of the priority ranking. The remainder of this section describes the mechanisms underlying these patterns.

7.2 Underlying Mechanisms

The are two reasons why a student’s assignment utility may change from one scenario to the other: either the student is administratively assigned in one scenario while he manages to be assigned to one of his listed choices in the other; or the student is assigned to one of his listed choices under both scenarios, but the very program he is assigned to differs between the two scenarios. Table A.6 in Appendix A shows the share of students affected by each margin of change, as well as the average welfare change associated with each margin. Under the single-phase procedure, 6.8% of students end up being administratively assigned. 80% of these students gain assignment under a two-phase implementation; more than 93% of them gain assignment under a five-phase implementation. Students who manage to get assigned to one of their choices because of the sequencing are those who, individually, experience the largest expected post-utility gains. The average expected utility gain associated with gaining assignment is about 1.9 to 2.7 times the overall average expected post-utility gain. While the average expected post-utility gain experienced by students who would be assigned under a single-phase implementation is lower (.94 to 1.04 times the overall average expected utility gain), they are still very large, equivalent to a 50 to 67km-reduction in university-home distance relative to the single-phase procedure.

Figure 11 is a first step towards bringing together the findings from Figure 10 and Table A.6. Figure 11 plots, as a function of priority ranking, the probability that a student fails to be assigned to any of her listed choices (and hence the probability that she ends up being administratively assigned). The round-mark plot represents the probability of administrative assignment under the single-phase benchmark. Two facts are
Figure 10: Expected change in *ex-post* utility as a function of priority, relative to one-phase benchmark

![Figure 10](image1.png)

The LHS panel shows expected welfare gains as a function of student priority ranking, expressed as a reduction in the km-distance between the student’s home region and his one-phase assignment. The RHS panel shows confidence intervals.

Figure 11: Probability of non-assignment as a function of priority

![Figure 11](image2.png)

The LHS panel shows students’ expected probability to be rejected from all their listed choices, as a function of their priority ranking, under the different implementation scenarios. The RHS panel shows confidence intervals. Read: Under the one-phase implementation, the student ranked 2000th in the priority ranking is unassigned with probability (close to) 0.

worth noting about this plot. First, the probability of administrative assignment increases from the top to the bottom of the priority ranking, that is, as the state of the world students need to forecast gets further from the one they have information on. Under the single-phase scenario, information about the choice set is provided once, before anyone submits their application list. Hence, the lower a student’s priority, the larger the number of students to be assigned before her. In other words, the larger the number of random events (assignments) to alter the initial state of the world before she gets to be assigned. Comparing the one-phase and the two-phase plots shows that the revelation of information, as done in the sequential Tunisian mechanism, increases low-priority students’ average expected *ex post* utility by bringing them “closer” to up-to-date information. Access to more accurate information about their choice set allows them to form an application list that is more likely to yield a match.

Second, the probability of administrative assignment under the single-phase benchmark is very close to zero for students in the top half of the priority ranking. In fact, virtually all of the extensive-margins gains accrue to students in the bottom 40% of the priority ranking. This means that, while Figure 10 shows utility gains for students of all priority levels, distinct mechanisms drive the gains experienced by students with high vs low priority.
Figure 12: Average selectivity of listed choices as a function of priority

Note: The black and gray plots in the LHS panel show the selectivity level of students’ first-listed choice as a function of their priority ranking, under the different implementation scenarios. As a reference, the red solid line shows the average selectivity level of students’ assignment in the perfect-information setting. The RHS panel shows confidence intervals. Read: Under the one-phase implementation, the student ranked 2000\textsuperscript{th} lists a first choice with average selectivity .68, that is, which filled up with a marginal student ranked at the 68\textsuperscript{th} percentile of the priority distribution in the previous year.

Figure 12 illustrates the mechanism underlying high-priority students’ ex post-utility gains. The black and gray dashed plots in the left (resp. right) panel show, for various implementation scenarios and as a function of priority, the average selectivity level (in terms of past-year cutoff) of students’ first-listed choice. As a reference, the red solid line shows the average selectivity level of students’ assignment in the perfect-information setting.

As compared to the single-phase benchmark, the two-phase implementation induces students in the first half of the priority ranking to dramatically increase the selectivity level of their listed choices, particularly for students in the second quarter of the priority ranking. This demonstrates the second channel through which the Tunisian mechanism affects students’ application behavior: the improvement of students’ fallback option if they fail to be assigned to any of their listed choices. Under a sequential implementation, students in earlier groups are given the chance to reapply with top priority in later groups if they fail to be assigned to their first set of listed choices. The division of the cohort in two application groups therefore dramatically increases the value of top-priority students’ fallback option. In the single-phase benchmark, it is equal to the expected utility of a program with leftover seats after the whole cohort submitted their application and got a chance to be assigned. Under the two-phase scenario it is equal to the expected utility of their favorite program among those leftover seats before anyone in the bottom half of the priority ranking is assigned. The lower a student’s priority in the top half, the smaller the difference between the set of programs expected to still have vacancies by the time she gets considered for assignment by the algorithm with her first-time application list, and the set of programs expected to still have vacancies by the time Group-2 students apply. In other words, the lower a student’s priority in the top half, the lower the ex post utility she expected to lose by having to reapply at the top of Group 2 instead of matching with a program the first time she applies. This increase in $V_i(0)$ in Equation (1) leads students students to increase the selectivity of their listed choices, throughout their whole application list —as an example, Figure 12 shows the 1st choice; Figure A.8 in Appendix A shows the 6th choice; other listed choices show similar patterns.

Note that Figure 12 also illustrates how the information-revelation mechanism, already discussed with Table A.6, generates utility gains at the bottom of the priority ranking. The selectivity of low-priority students’ listed choices is much higher under the one-phase implementation than it is under multiple-phase implementations. This is because under a multiple-phase implementation, it has been revealed to low-priority students
that the programs they would have applied to in the one-phase procedure have filled up. The high selectivity of low-priority students’ listed choices in the one-phase scenario, and hence the large decrease in selectivity of their choice caused by information provision, are driven by their expectations: Section 6 estimates show that they tend to expect cutoffs to decrease from one year to the next.

Figure 13: Expected change in assigned utility: with and without reapplication chance

The plots are analogous to those in Figure 10, but in scenarios in which the fallback option channel is shut down. The LHS panel shows expected welfare gains as a function of student priority ranking. The RHS panel shows confidence intervals. See the note under Figure 10 for details.

To disentangle the gains from the information-revelation and fallback-option channels, Figures 13 and 14 show results analogous to Figures 10 and 9 but for multi-phase scenarios in which the fallback-option mechanism has been shut down. That is, I simulate students’ choices in implementations in which students are given updated information about vacancies at the beginning of each phase but are told they would be administratively assigned at the very end of the procedure if they fail to be assigned to their listed choices. As a reference, Figure 13 reproduces plots from Figure 10: the solid red plot shows the difference in expected student welfare between the perfect-information setting and the single-phase benchmark; the thinner black and gray plots (the latter is mostly confounded with the red plot) show welfare gains from an implementation in two and five phases, respectively. Now, the thick black and gray plots show welfare gains from information-only scenarios —with two and five phases, respectively. When no second chance is offered, students in the first group of a multi-phase implementation are facing exactly the same choice problem as in the single-phase benchmark. Hence, as expected, these students’ expected assignment utility is the same under the benchmark and an information-only scenario. Welfare gains discontinuously jump at the beginning of the second group of the multi-phase scenario, reaching even higher levels than under the corresponding multi-phase scenarios with second-chance. This sheds light on the fourth pattern initially documented by Figure 10. High-utility seats not taken by high-priority students who “under-place” in the one-phase procedure (or, here, in the absence of the fallback-option channel) are seats that remain available to Group-2 students, while they would not have been in the perfect-information setting or in a multi-phase scenario with fallback-option. These “general equilibrium” effects are the reason why increasing the number of phases, or even moving from a small number of phases to perfect information, is not Pareto improving. The comparison of the two-phase and perfect-information scenarios in Figure 10 illustrates this strikingly: when moving from two phases to perfect information, part of the welfare gains going to students in the first half of the priority ranking are taken from students ranking at the beginning of the second half, who actually experience a decrease in utility.

Figure 14 decomposes total student welfare gains shown in Figure 9 into gains from the information-revelation channel, and gains from the fallback-option channel. Relative to a single-phase implementation,
a two-phase scenario with no fallback-option procedure increase average utility by the equivalent of a 28km-reduction in students’ distance to their single-phase assignment. This is only 58% of the average ex post utility gain achieved with a two-phase scenario with both information revelation and a second-chance procedure, as is done in Tunisia. Figure 14 shows that as the number of phases increases, an information-channel-only procedure can generate an increasingly large share of the welfare obtained with a multiple-phase scenario with both information revelation and a second-chance procedure: a three (resp. five)-phase procedure with no fallback-option procedure achieves 76% (resp. 91%) of the utility gains obtained with a three (resp. five)-phase scenario allowing for both channels.

Figure 14: Expected change in average utility: effect of information vs reapplication chance

The black plot is the same as in Figure 9. The blue plot is analogous but shows the share of potential welfare gains $\Delta W_{PI,1}$ that can be achieved with information-only scenarios. Read: Using a two-phase procedure with no fall-back option, instead of the standard single-phase implementation, allows to realize 49% of the welfare gains that would be achieved by moving from one phase to perfect information.

### 7.3 Discussion

Kapor, Neilson, and Zimmerman (2018) have noted the importance of accounting for students’ subjective expectations about their admission chances to assess the welfare consequences of assigning students using a manipulable mechanism like the Boston mechanism instead of a strategy-proof mechanism. Here also, the magnitude of gains from sequencing depend on applicants’ subjective expectations about their admission chances. Estimates from the expectations-formation model show that low-priority students behave as if their admission chances to all programs were higher than they actually are — as if the selectivity of all programs would decrease relative to the previous year. This amplifies their propensity to list programs that, ex post, turn out to be full. As a consequence, the revelation of information significantly changes their application choices and, in turn, their assignments. Estimates from the expectations-formation model show that, on the contrary, high-priority students tend to behave as if their admission chances to all programs were lower than they actually are — as if the selectivity of programs would increase relative to the previous year. This magnifies their tendency to apply to low-selectivity program in a one-phase setting. The provision of a higher-value fallback option therefore changes all the more their application choices.

It is worth emphasizing that the crucial effect of the value of students’ fallback option on their application choices and, therefore, their assignments has important consequences in any school choice context if all students do not have equal access to high-quality fallback options. It implies, for instance, that in a school district using a assignment procedure that is not strategy-proof, a student whose family can afford a good private school is more likely to apply to the best local public schools than a student with same preferences and same priority whose fallback option is a lower-quality public school. Akbarpour and van Dijk (2018) discuss this mechanism and its consequences in a theoretical framework. Providing empirical evidence of this
mechanism, though, requires knowledge of the value of students’ fallback option. When the fallback option is an “outside” option (that is, not available within the centralized assignment mechanism, like a private school in the context of public school choice), individual data about fallback options is typically limited. The Tunisian context allows to provide such empirical evidence, as students’ fallback option is built-in the assignment mechanism.

8 Conclusion
Strategy-proofness and assurance of a fair matching are qualities that policy-makers greatly value in school choice mechanisms. While these qualities are theoretical properties of the unrestricted-list DA, list-size restrictions are a feature of virtually every implementation of the DA, and it has proven to be hard to go without them in practice. This paper shows that an easy-to-implement variant of the restricted-list DA can mitigate uncertainty and yield outcomes very close to what would obtain under the unrestricted-list DA.

Quantifying the consequences of list-size restrictions on students’ applications, assignments, and ex post utility, and assessing the extent to which a change in the design of assignment mechanisms can mitigate them require knowledge of both students’ preferences for schools and their expectations about their admission chances. An interesting implementation of the restricted-list DA in Tunisia allows me to circumvent the challenge of identifying both preferences and expectations using administrative data. I argue that the Tunisian design induces a subset of students to truthfully report their most-preferred programs, allowing me to recover students’ preferences for post-secondary programs. In a second step, I characterize students’ expectations about their admission chances as those rationalizing other students’ observed application lists, given identified preferences.

My empirical analysis builds on a particular version of the DA in which schools all use the same priority ranking to determine admissions—the serial dictatorship. I leave it to future work to extend the model to a more general version of the DA allowing for school-specific priorities. Beyond school choice and the DA, an interesting avenue for future research is to understand how a sequential implementation could improve the outcomes of mechanisms used in other settings, such as professional matches or public housing.

References


