Need vs. Merit: The Large Core of College Admissions Markets

Avinatan Hassidim
Bar-Ilan University

Assaf Romm
Hebrew University of Jerusalem

Ran I. Shorrer
Penn State University

November 29, 2017

Abstract

We study college admissions markets, where each college offers multiple levels of financial aid. Colleges subject to budget and capacity constraints wish to recruit the best qualified students. Deferred Acceptance is strategy-proof for students, but the scope for manipulation by colleges is substantial, even in large markets. Successful manipulation takes the simple form of allocating funding based on need rather than merit. Stable allocations may differ in the number of assigned students. In Hungary, where the centralized college admissions clearinghouse uses Deferred Acceptance, another stable allocation would increase the number of students accepted to college by at least 3%, and applicants from low socioeconomic backgrounds would benefit disproportionately.

1 Introduction

In recent years, a growing number of students are being assigned to schools through centralized clearinghouses. The success of such clearinghouses crucially relies on the use of a stable matching mechanism.[1] Stability is also

[1]See Roth and Xing (1994) and Roth (2002).

useful for predicting behavior in decentralized matching markets.[2] Empirical and theoretical studies suggest that all applicants, save for a handful, receive the same assignment in all stable allocations.[3] This finding, that the *core*, the collection of stable allocations, is small has several implications. First, a designer who wishes to implement a stable outcome has limited scope for further design (e.g., affirmative action). Second, incentives to misreport one's preferences to the student-proposing Deferred Acceptance mechanism (henceforth DA[4]) are minimal (Azevedo and Budish, 2012).[5]

The above-mentioned results apply to settings where agents on each side of the market (e.g., students and schools, men and women, etc.) have preferences over potential partners from the other side. However, the environments studied and designed by economists are often more complex. For example, in college admissions markets, universities often offer admission to to several study programs and multiple levels of financial aid. These more complex environments are studied in the matching-with-contracts literature (Hatfield and Milgrom, 2005). Much of this literature focuses on identifying conditions under which DA remains stable and strategy-proof for students. The motivating question of this paper is: *Do the findings on the size of the core and the good incentives properties generalize in the presence of contracts?*

They do not. We study a natural extension of the large (matching-without-contracts) market model of Kojima and Pathak (2009) that captures the structure of preferences observed in centralized college admissions markets.[6] We observe that under DA, financial aid decisions are based on merit

[2]See Banerjee et al. (2013).

[3]A large portion of the literature on the theory of two-sided matching markets is motivated by the potential multiplicity of stable allocations. Examples include studies of the structure of the core (the collection of stable allocations) (Knuth, 1976), of fair stable matchings (Teo and Sethuraman, 1998; Klaus and Klijn, 2006; Schwarz and Yenmez, 2011), of the extent to which it is possible to improve the allocation of under-demanded hospitals or to increase the number of assigned doctors (Roth, 1986), and of incentives (Roth, 1982; Sönmez, 1999; Ehlers and Massó, 2007).

[4]Unless mentioned specifically, DA refers to the student-proposing version of DA (Gale and Shapley, 1962).

[5]Truthful reporting to the student-proposing DA mechanism is a weakly dominant strategy for students, and there is no stable matching mechanism that makes truthful reporting dominant for both sides of the market (Dubins and Freedman, 1981; Roth, 1982). Numerous studies analyzed the optimal behavior of schools when the student-proposing DA mechanism is in place (examples include Sönmez, 1997; Roth and Rothblum, 1999; Ehlers, 2004; Konishi and Ünver, 2006; Coles, Gonczarowski and Shorrer, 2014).

[6]The theory of matching with contracts has many other applications, such as the al-

(Theorem 1),[7] and thus ignore students' outside alternatives. Based on this observation, we show that the expected fraction of students and colleges that have multiple stable allocations does not vanish as the the size of the market grows large, and the same is true for the fraction of colleges that can successfully manipulate DA. Furthermore, the manipulation we identify takes a simple form that can be interpreted as offering need-based – rather than merit-based – financial aid.

Empirically, we corroborate the predictions of the model using data from two centralized college admissions markets, where variants of the student-proposing DA are in use. We show, based on reported preferences, that thousands of students have multiple core allocations. Moreover, we show that the number of students accepted to colleges in Hungary could increase by more than 3% by switching from the DA outcome to another stable allocation, and that colleges could substantially improve the quality of their incoming cohorts by misrepresenting their preferences.

We next provide two examples illustrating the key ideas behind our main results. The first example shows that stability is not compromised if a college refuses financial aid to a student whose choice of college is not sensitive to the availability of financial aid. The second example shows that by refusing funding to such students, the college's budget constraint is relaxed, which may allow it to recruit better, price-sensitive students. Our main theoretical result (Theorem 2) shows that there are many instances in which colleges can manipulate DA by reporting that price-insensitive students are ineligible for admission with financial aid, and that this manipulation results in a stable allocation that these colleges prefer. Intuitively, the reason is that under DA financial aid decisions are based on merit, and thus ignore students' outside alternatives. Readers who find this verbal description satisfactory may want to skip the examples and go directly to Section 1.2.

1.1 Examples

The following examples use the notation of the paper, but since the model has not yet been introduced, we provide a verbal description for each piece

location of cadets to military branches (Sönmez, 2013; Sönmez and Switzer, 2013) and entry-level labor markets (Niederle, 2007; Dimakopoulos and Heller, 2015). While the focus of this paper is on college admissions markets, our results extend to these environments.

[7] We are grateful to Joel Sobel for this observation.

of notation.

Example 1. There are n colleges and m students. Colleges offer positions with or without financial aid (formally, $T = \{0, 1\}$). Each college has a capacity of one and prefers all student and financial aid combinations to the outside option of staying unmatched (formally, $q_0 = q_1 = 1$). For all colleges, accepting a student s without financial aid is preferred to accepting the same student with financial aid. However, colleges care lexicographically more about the identity of the student than about funding (i.e., if accepting s without financial aid is preferred to accepting s' with financial aid, then accepting s with financial aid is also preferred to accepting s' with financial aid). Students prefer admission with financial aid, but they too care lexicographically more about the identity of the college than about funding. Students also prefer any allocation to staying unmatched.

Claim 1. *Under the above conditions, there are* $\min\{m, n\}$ *colleges and* $\min\{m, n\}$ *students with multiple stable allocations.*

Proof. Since colleges have strict preferences and unit demand, there exists a stable matching and the same agents are matched in all core allocations (Hatfield and Milgrom, 2005). Since all agents find any allocation acceptable, stability implies that there cannot be unmatched agents on both sides of the market. Hence, $\min\{m, n\}$ colleges and students have some stable allocation other than being unmatched. Given a stable allocation, changing the terms while maintaining the identities of the contracting parties maintains stability. □

Note that even if colleges know nothing more than the above description about others' preferences, they have a simple manipulation for the student-proposing version of DA: by declaring all funded contracts unacceptable and maintaining the rest of their strategy, they will be assigned the same students, but will not have to provide financial aid.[8] The driving force behind Example 1 is the near indifference of *both* sides of the market to funding. The requirement that all agents be acceptable does not play an important role: in a model without this assumption, the only difference is that $\min\{m, n\}$ would be replaced with the cardinality of some stable matching.

[8]Under the assumption that students will use their dominant strategy of reporting their preferences truthfully.

While the market described in Example 1 has a large core in the sense that many agents have multiple stable allocations, the proof relies on allocations in which the same parties contract with each other, and only the contractual terms differ. Given that we assumed that funding is not particularly important for any of the participants (relative to the identity of their partner), the core may still be small in the sense that the same agents are matched in all stable allocations and that agents do not have "strong" preferences between stable allocations.

In the next example students differ in the importance they attribute to financial aid, and the single college can accept more students than it can fund. The example highlights several differences between our college admissions environment and the one studied by Gale and Shapley (1962), where each college offers only one bundle of contractual terms. Notably, in our environment there is no student-optimal stable allocation, and the number of students attending college differs between stable allocations.

Example 2. There is one college ($C = \{h\}$) with two seats, but only one scholarship available ($T = \{0, 1\}$, $q_h^0 = 2$, $q_h^1 = 1$), and two students, $S = \{r, p\}$. It may help to think of r as a "rich" applicant and p as a "poor" applicant. The rich applicant, r, also happens to be a better "fit" with college h ($r \gg_h p$). The college prefers to accept the best "fit" students, and to fill its capacity. The college's preferences over acceptable feasible allocations are summarized as follows:

$$\{(r,h,0),(p,h,0)\} \succ_h \{(r,h,1),(p,h,0)\} \succ_h \{(r,h,0),(p,h,1)\} \succ_h$$

$$\succ_h \{(r,h,0)\} \succ_h \{(r,h,1)\} \succ_h \{(p,h,0)\} \succ_h \{(p,h,1)\} \succ_h \emptyset,$$

where 1 (0) indicates admission with(out) financial aid. Applicant r's preferences are $(r,h,1) \succ_r (r,h,0) \succ_r \emptyset$. That is, she would prefer to receive financial aid, but is willing to attend h even if she does not get it. The poor applicant, p, is only interested in admission with financial aid. Thus, his preferences are summarized by $(p,h,1) \succ_p \emptyset$. Under these preferences, there are two stable allocations, $\{(r,h,1)\}$, which is the result of the student-proposing DA algorithm, and $\{(r,h,0),(p,h,1)\}$.

Notably, the two allocations have different numbers of assigned students. Moreover, the outcome of DA is not the stable allocation most preferred by both students. In fact, such an allocation does not exist.[9]

[9]To see that there is no stable outcome that is most preferred by both students, note

1.2 Overview of the results

Our model of college admissions markets is a natural extension of the Kojima and Pathak (2009) model of large matching markets (without contracts). In their model, while the number of schools is large, each student finds a small number of them acceptable. We augment their model by introducing different levels of financial aid in the same college. We assume that colleges face a constraint on the number of financial aid packages at each level, but otherwise have no strong preferences over the amount of financial aid they provide and over the identities of the funded students in a given cohort. We also assume that whenever an applicant finds a certain college acceptable under some terms, she prefers to attend that college under a more generous financial aid package.

Our assumptions on the demand structure reflect many features of the preferences reported to two centralized college-admissions matching-with-contracts markets: the Israeli Psychology Master's Match (IPMM; Hassidim, Romm and Shorrer, 2017) and Hungarian college admissions (Biró, 2007).[10] The assumption on colleges' preferences reflects the reports of departments in the IPMM, and is consistent with the choice functions used in the Hungarian market.

We prove that the expected fractions of applicants and of colleges with multiple core allocations are large (non-vanishing in large markets), and that the same holds for the fraction of colleges that can successfully manipulate DA when all other agents are truthful. Furthermore, different core allocations may result in substantially different numbers of students admitted to college.

We corroborate our theoretical predictions using administrative data from both the IPMM and the Hungarian market. We find that in Hungary, the number of students admitted to college under the current student-proposing algorithm (about 60,000 a year) could be increased by more than 3% by choosing a different core allocation. Applicants from a lower socioeconomic background would benefit disproportionately,[11] as would female and rural ap-

that the only allocation that both students weakly prefer to both core allocations has both of them receiving financial aid, but this allocation is not acceptable to the college.

[10]The IPMM is strategy-proof for applicants. Like most real-life implementations of DA, the Hungarian college admissions mechanism is not, strictly speaking, strategy-proof (Pathak and Sönmez, 2008; Shorrer and Sóvágó, 2017). Outside of the matching literature, Avery and Hoxby (2004) make similar assumptions on students' preferences. They too find empirically that students apply to a limited number of colleges.

[11]Intuitively, the reason is that under DA financial aid is allocated based on merit

plicants. Moreover, colleges could successfully manipulate DA and "poach" talent from their competitors. Our findings stand in sharp contrast with those of Roth and Peranson (1999) who find that in the early 1990s, only about 0.1% of approximately 20,000 applicants to the NRMP had multiple stable allocations.

To provide intuition, we highlight the main differences between our model and that of Kojima and Pathak (2009). The driving force behind the small core result of Kojima and Pathak is what they call the *"vanishing market power"* of schools when the student-proposing version of DA is used. Namely, a school's strategic rejection of a student is highly unlikely to result in a proposal from another, preferred student, a necessary condition for the existence of a profitable manipulation in their setting.

By contrast, the driving force behind the large core in our model is the presence of *market power*. Given a core allocation, colleges have market power over students who receive financial aid, but have no outside option that they prefer to a contract with the same college at a lower level of financial aid. An extreme case is when an applicant ranks all contracts with the same college consecutively. Colleges can "price discriminate" by not offering funding to such students (as illustrated in Example 1).[12] And, the freed-up funds may be used to recruit price-sensitive students.

The presence of market power in our model is the result of colleges offering several contracts that the same individuals tend to be interested in. In contrast with the environment without contracts, colleges can successfully manipulate DA without generating offers from additional students. Instead, it is sufficient that a strategically rejected student apply for admission with lower financial aid, as this would relax the college's budget constraint, and allow it to accept a price-sensitive student it would have otherwise had to reject. This rationale is illustrated in Example 2, where the college "loses" in the funded seat, by accepting the lower-quality student, so as to gain overall through an improved assignment (of the student who would have received

(as illustrated by Example 2). This intuition extends to other mechanisms, such as the contract-proposing version of DA, where each contract is regarded as a separate program and the program-proposing variant of DA is used. This mechanism is used in the Turkish centralized college-admissions clearinghouse (Balinski and Sönmez, 1999). A World Bank report questions the current allocation system and states that it "is akin to giving a large number of scholarships in each institution on the basis of merit" (Hatakenaka, 2006).

[12]Economics Ph.D. programs typically consider prospective students' outside options and often offer to "match" offers from other programs.

funding under truthful reporting) to the unfunded seat.

Our results shed light on the policy debate around market power in higher education (e.g., Hoxby, 2000). This literature gained traction after the Department of Justice pursued an antitrust case against a group of elite colleges for sharing prospective students' financial information and coordinating their financial aid policy. MIT contested the charges, claiming that this practice prevents bidding wars over the best students and thus frees up funds to support needy students, and that the school did not profit financially from this practice (DePalma, 1992). In 1994, Congress passed the Improving America's Schools Act, whose section 568 permits some coordination and the sharing of information between institutions with a need-blind admissions policy. Our model provides theoretical support to MIT's arguments. We show that even in the absence of a motive to increase profit, schools have an incentive to apply market power in order to improve the quality of their incoming cohorts, and that the consequences for students are heterogeneous. In particular, the direct effect is that some (needy) students gain and other (wealthy) students lose. Furthermore, the model illustrates how information about students' alternatives can facilitate such behavior.

Finally, we make two smaller contributions, which we expand on in the related literature section. First, we show that, in contrast with the environment studied by Gale and Shapley (1962), a college may have multiple core allocations but not be able to manipulate DA. Second, we show that while DA always terminates in a core element of the college admissions market, the class of college admissions markets cannot be embedded, in the sense of Echenique (2012), into Kelso and Crawford (1982) matching-with-salaries markets.

1.3 Related literature

Our study is most closely related to papers studying the size of the core in two-sided matching markets. There are many ways to think about the size of the core, and thus multiple notions of "smallness." One such notion is that the same agents are matched in all elements of the core. This result is part of the rural hospital theorem proved by Roth (1984*a*, 1986) for the case of many-to-one markets with responsive preferences, and extended by Hatfield and Milgrom (2005) to environments with contracts where programs' preferences meet certain conditions.

An alternative notion of smallness is the number of agents who receive a different assignment in different core allocations. Early studies of the size of the core focus on one-to-one markets with the same number of men and women, where all members of the opposite sex are acceptable. In a random instance of such a market where preferences are drawn independently and uniformly at random, the core is typically large in this sense (Pittel, 1989, 1992).

Roth and Peranson (1999) show that the core in the National Residency Match Program (NRMP) is small in this sense.[13] Roth and Peranson (1999) attribute their finding to the market being large and students ranking only a small number of residency programs, and provide simulation evidence in support of their theory. This finding is later proved theoretically by Immorlica and Mahdian (2005), Kojima and Pathak (2009), and Storms (2013), in increasingly general environments. Under an additional regularity condition, Kojima and Pathak (2009) prove that truthful reporting to DA is an approximate Bayesian Nash equilibrium.[14] Ashlagi, Kanoria and Leshno (2017) show that the core is typically small even when all members of the opposite sex are acceptable, as long as the number of men and women is not exactly equal. Azevedo and Leshno (2016) study a model with a finite number of schools and a continuum of students, and find that the core is generically unique.

Another notion of smallness is that the difference in utility between different core allocations is small for all (or most) agents. Holzman and Samet (2014) find that the core is small in this sense when preferences are correlated. Lee (2017) allows for correlation in preferences through a common-value component and finds that under certain conditions on preferences the core is small, and so are incentives to misreport under DA in one-to-one markets.[15] In Section 3 we show that the core of large college admissions markets is large relative to all three of the above-mentioned senses.

Our paper is related to the growing literature on matching with contracts (e.g., Kelso and Crawford, 1982; Roth, 1984*b*; Fleiner, 2003; Hatfield and

[13] Hitsch, Hortaçsu and Ariely (2010) document a small core in an online dating market, and Banerjee et al. (2013) find similar results in Indian marriage markets. Menzel (2015) shows that in large markets the core converges in the sense that the probability of a man of a certain type to be matched with a woman of a certain type converges to a unique limit.

[14] Azevedo and Budish (2012) show that DA is strategy-proof in the large.

[15] Coles and Shorrer (2014) show that even under incomplete information the exact best response of schools under DA can be substantially different from truthful reporting.

Milgrom, 2005; Hatfield and Kojima, 2010; Hatfield and Kominers, 2015; Hatfield, Kominers and Westkamp, 2015). Several authors study college admissions environments using the matching-with-contracts model (Abizada, 2016; Afacan, 2017; Aygün and Bó, 2016; Pakzad-Hurson, 2014; Westkamp, 2013; Yenmez, 2015), but they focus on other questions.

Closely related are studies of reserve design in the context of school choice (Dur et al., 2013; Dur, Pathak and Sönmez, 2016). A key observation in this literature is that applicants are indifferent between different seats in the same school, which implies, using our terminology, that schools have market power over all assigned students. The reserve-design literature focuses on the effect of different ways the mechanism can break these preference ties to form strict student rankings of contracts, while keeping the priorities at each seat fixed. By contrast, we study an environment where students have strict preferences over all contracts, and concentrate on changes to colleges' preferences.

In the one-to-one setting, a school has an incentive to misreport its preferences to the student-proposing DA mechanism if and only if the school has multiple core allocations (Demange, Gale and Sotomayor, 1986). In many-to-one markets, only one implication is correct (e.g., Kojima and Pathak, 2009). Namely, given a profile of preferences, a school may have a unique stable allocation and still have an incentive to misrepresent its preferences to a DA mechanism. We show that in college admissions markets neither statement implies the other. Namely, a college may also have multiple core allocations but no incentive to misrepresent its preferences to a DA mechanism.

Echenique (2012) shows that when colleges' preferences satisfy the substitutes condition of Hatfield and Milgrom (2005), there exists an embedding that maps markets with contracts into Kelso and Crawford (1982) markets with salaries and gross substitutes demands, such that the set of stable allocations is preserved.[16] The construction does not apply to environments like the one studied by Sönmez and Switzer (2013), where only the weaker condition of unilateral substitutability (Hatfield and Kojima, 2010) is satisfied. Schlegel (2015) and Jagadeesan (2016) construct different embeddings that extend to this broader class of markets. We contribute to this literature by identifying a real-life environment where DA is stable and strategy-proof for students, but that cannot be embedded into a Kelso and Crawford environment. To see this, note that the college admissions market in Example 2 does not have a student-optimal stable allocation, a necessary condition for

[16]Kominers (2012) extends this result to the many-to-many setting.

the existence of an embedding.

The paper is organized as follows. Section 2 presents our model and discusses the differences between our college admissions environment and the environment studied in Gale and Shapley (1962). Section 3 presents the main theoretical result, and a proof of a special case that illustrates key ideas behind the proof. The complete proof is in Appendix B. Section 4 presents the empirical evidence. Section 5 discusses the implications of our findings and proposes directions for future research.

2 Model

We use the many-to-one matching-with-contracts model of Hatfield and Milgrom to describe college admissions environments. There is a finite set of *colleges*, C, a finite set of *students*, S, and a finite set of *contractual terms*, T. A *contract* is a tuple $(s, c, t) \in S \times C \times T$ that specifies a student, a school, and the contractual terms that govern their relationship. In this paper, $t \in T$ will typically describe the level of financial aid. The set of all possible contracts, X, is a subset of $S \times C \times T$.

We denote by X_i the set of all possible contracts that involve agent $i \in S \cup C$. Each agent, i, has strict preferences over subsets of X_i, which we denote by $\succ_i$. We follow the convention of (often) omitting sets that are ranked lower than the empty set from the description of $\succ_i$.

An *allocation* is a subset Y of X. Given an allocation Y, we sometimes refer to $Y_i := Y \cap X_i$ as agent i's allocation. An allocation Y is *individually rational* if for any agent i the entire set Y_i is the most preferred subset of Y_i.

Finally, we assume that all students prefer the empty set to any subset with cardinality strictly greater than 1. Given that our interest is in individually rational allocations only, this encodes our assumption that the market is a many-to-one matching market. An allocation Y is *feasible* if $|Y_s| \leq 1$ for each student s.

Financial aid

Unless specifically mentioned, we assume that T is a finite subset of $\mathbb{N}$. We think of $t \in T$ as a funding level, and assume that for each student, s, and college, c, $(s, c, t) \succ_s (s, c, t')$ if and only if $t > t'$. As we discuss later, this modeling choice is not crucial for any of our results, but we make it in order to capture key aspects of our datasets.

We make several assumptions on colleges' preferences. First, each college, c, is associated with a sequence of numbers, $\{q_t^c\}_{t\in T}$, such that if $t < t'$ then $q_t^c \geq q_{t'}^c$. The number q_t^c represents a constraint on the number of students that can be accepted with funding level t or higher. Each college, c, prefers the empty allocation to all allocations that *violate c's quotas*, that is, assign to c more than q_t^c students with funding level t or higher.[17]

Second, each college, c, has a *master list*, a complete ranking over $S\cup\{\emptyset\}$, denoted by $\gg_c$. Given an allocation Y, a contract $(s, c, t) \in Y_c$, and a contract $(s', c, t) \notin Y_c$, $Y_c \succ_c Y_c \setminus (s, c, t) \cup (s', c, t)$ if and only if $s \gg_c s'$. Moreover, if Y_c does not violate c's quotas, $Y_c \succ_c Y_c \setminus (s, c, t)$ if and only if $s \gg_c \emptyset$. In words, the college will accept an additional contract as long as the student is ranked higher than the empty set on the college's master list, and the contractual terms will not cause a quota violation.

Finally, we assume that, as long as quotas are not exceeded, the allocation of funding is less important than the identities of incoming students. Formally, if $Y_c \succ_c \emptyset$ and $Y'_c \succ_c \emptyset$, and there exists a bijection $\phi : Y_c \to Y'_c$ such that $x \in X_s \iff \phi(x) \in X_s$ for all $x \in Y_c$ and all $s \in S$ (i.e., Y_c and Y'_c differ only in contractual terms), but no such bijection between Y_c and Z_c (roughly, Y_c and Z_c differ in contracting parties), then $Z_c \succ_c Y_c \iff Z_c \succ_c Y'_c$. This assumption is realistic insofar as funding terms and quotas are often exogenously given.[18]

The special case of our model where $|T| = 1$ corresponds to two-sided many-to-one matching-without-contracts markets with responsive preferences, the environment studied by (Kojima and Pathak, 2009).

[17] In the language of NRMP, we allow for "reversions": donating unfilled positions in one program to another (Roth and Peranson, 1999; Niederle, 2007). All of our results hold in an alternative model in which each college has a constant capacity for each level of funding, but seats at one level cannot be converted to seats at other level.

[18] For example, funding may come from the government in the form of a full tuition waiver (see Artemov, Che and He, 2017). Alternatively, the "college" in our model may represent a department that is free to make admissions decisions, but whose funding policy is set by the institution and whose budget is earmarked (see Hassidim, Romm and Shorrer, 2017).

Choice functions and stability

Agent i's preferences induce a *choice function*, $Ch_i : 2^X \to 2^{X_i}$, that identifies the subset of Y_i most preferred by i from any subset Y of X. Formally, $Ch_i(Y) := \max_{\succ_i} Z \subset Y_i$.[19] An allocation Y is *unblocked* if there does not exist a college, c, and $Z \subset X_c \setminus Y$ such that $Z_i \subset Ch_i(Z_i \cup Y_i)$ for all $i \in S \cup C$. An allocation Y is *stable* (alternatively, in the *core*), if it is individually rational and unblocked.

A remark on our choice to make assumptions directly on preferences, rather than on choice functions, is in order. It is well known that, in general, assumptions on choice functions can be less restrictive than assumptions on preferences. We choose to make assumptions directly on preferences for two reasons. First, we think that this makes the assumptions more transparent. Second, these assumptions reflect our understanding of colleges' preferences based on our practical experience (e.g., Hassidim, Romm and Shorrer, 2017).

2.1 Properties of college admissions environments

We begin this section by noting that we did not restrict colleges' choice functions to feasible allocations (i.e., at most one contract per student). Upon making this observation, it is easy to verify that the choice functions satisfy the hidden substitutes condition of Hatfield and Kominers (2015), and that they meet the other conditions of their Theorems 1–3 that assure that DA yields a stable allocation and is strategy-proof for students.

Proposition 1. *The core of college admissions environments is non-empty. Furthermore, DA terminates in a core element, and the mechanism it induces is strategy-proof for students.*

Proposition 1 lists some similarities between our model and the matching-without-contracts environments studied by Gale and Shapley. Critically for our main result, it assures that the core is non-empty and that DA terminates in a core allocation.

The following theorem formalizes our claim that under DA financial aid is distributed based on merit.

[19] The fact that choice functions are derived from strict preferences implies that they satisfy the irrelevance of rejected contracts condition (Aygün and Sönmez, 2013).

Theorem 1. *Let* $\left\langle S, C, T, \{\succ_c, \gg_c, q_t^c\}_{t \in T}^{c \in C}, \{\succ_s\}_{s \in S}\right\rangle$ *be a college admissions market, and let* Y *be the stable allocation that corresponds to the outcome of DA. Then for all* s *and* s' *in* S*, for all* c *in* C*, and for all* t *and* t' *in* T*, if* (s, c, t) *and* (s', c, t') *belong to* Y*, then* $s \gg_c s'$ *implies* $t \geq t'$*.*

Proof. Omitted. □

We next highlight key differences between our model and the matching-without-contracts environments studied by Gale and Shapley.

Proposition 2. *A college admissions market may have no stable allocation that is most preferred by all students. Furthermore, different core allocations may have a different numbers of assigned students.*

Proof. Follows from Example 2 in Section 1.1. □

Corollary 1. The class of college admissions markets cannot be embedded (in the sense of Echenique, 2012) into Kelso-Crawford markets with salaries and gross substitutes demands.

The corollary follows by the fact that the existence of a student optimal stable allocation is a necessary condition for an embedding to exist.

Proposition 3. *i) In a college admissions market, a college may have multiple stable allocations and not be able to manipulate DA when all other agents are truthful. ii) Furthermore, it may have a unique stable allocation and be able to manipulate DA when all other agents are truthful.*

Proof. The second part is well known (e.g., Kojima and Pathak, 2009). The first part follows from Example 3. □

Example 3. We add to the (private) college h from Example 2 another (community) college, c, with one seat and no scholarships available ($C = \{h, c\}$, $q_0^c = 1$, $q_1^c = 0$). The poor student's first choice is the funded seat at h, and while he now prefers the unfunded seat at h to staying unmatched, he finds the (cheaper) community college more attractive. The rich student prefers college h to the community college under any funding terms.

Formally, students' preferences are now

$$(r, h, 1) \succ_r (r, h, 0) \succ_r (r, c, 1) \succ_r (r, c, 0) \succ_r \emptyset, \text{ and}$$

$$(p, h, 1) \succ_p (p, c, 1) \succ_p (p, c, 0) \succ_p (p, h, 0) \succ_p \emptyset,$$

c's preferences are

$$(r, c, 0) \succ_c (p, c, 0) \succ_c \emptyset,$$

and h's preferences remain

$$\{(r, h, 0), (p, h, 0)\} \succ_h \{(r, h, 1), (p, h, 0)\} \succ_h \{(r, h, 0), (p, h, 1)\} \succ_h$$

$$\succ_h \{(r, h, 0)\} \succ_h \{(r, h, 1)\} \succ_h \{(p, h, 0)\} \succ_h \{(p, h, 1)\} \succ_h \emptyset.$$

Under these preferences, there are two stable allocations: $\{(r, h, 1), (p, c, 0)\}$, which is the result of the student-proposing DA algorithm, and $\{(r, h, 0), (p, h, 1)\}$. The number of students attending each college is different between the two stable matchings. Again, the outcome of DA is not the stable allocation most preferred by all students. And the only allocation that both students weakly prefer to both of the above allocations requires both of them to be funded in the private college, which violates the college's individual rationality constraint. It is easy to verify that, when others are truthful, c cannot do better than the outcome of DA, as r must be assigned to h regardless of c's strategy.

The fact that c cannot manipulate DA even though it has multiple stable allocations is not due to one of them being the empty allocation. We could have augmented the example by adding a third student, d, that is only interested in the community college, and whom all colleges least prefer. The stable outcomes would be $\{(r, h, 1), (p, c, 0)\}$, which is the result of the student-proposing DA algorithm, and $\{(r, h, 0), (p, h, 1), (d, c, 0)\}$, but c would still have no incentive to misrepresent its preferences to DA when others are truthful.

The example also illustrates the role of outside options. The private college has market power over the rich student in the allocation that results from DA because she has no outside option that she prefers to the unfunded contract with the college. The college does not have market power over the poor student in the second stable allocation since, if he is not offered funding, he prefers to attend another school that will accept him. Example 4 in the appendix shows that given a stable allocation, a college may have market power over an assigned student who ranks other contracts between that student's allocation and another contract with the college, as long as the colleges who are parties to these contracts are not interested (and thus the student cannot form a blocking coalition with them). This observation proves useful for our empirical analysis.

It is worth pointing out that in the previous example, if the student-proposing version of DA is used, the private college has a simple profitable manipulation of declaring that the rich student is not eligible for financial aid. This manipulation is also feasible in a large market where the wealth level of applicants is known, and nothing else is known about students' preferences other than that financial aid does not alter rich applicants' preferences between colleges.

3 Theoretical Evidence

By now, it is clear that college admissions environments have some properties that distinguish them from markets without contracts. We now turn to address our main questions: *How likely can a college manipulate the student-proposing DA mechanism*, and *how likely does a college (student) have multiple core allocations?* To this end, we introduce the following random environment that generalizes the Kojima and Pathak model of large matching markets to college admissions environments.

3.1 Regular sequences of college admissions markets

Let a *uniform random market* be a tuple $\tilde{\Gamma} = \left\langle S, C, T, \{\succ_c, \gg_c, q_t^c\}_{t\in T}^{c\in C}, k\right\rangle$, where k is an integer greater than one, and $\succ_c$ represents college c's strict preferences (which must be consistent with its master list, $\gg_c$, and list of quotas, $\{q_t^c\}_{t\in T}$). A uniform random market induces a college admissions market by drawing students' preferences randomly in the following way:

- Step 1: for each student independently, draw k different colleges from C uniformly at random.
- Step 2: for each student independently, draw uniformly at random an acceptable permutation over the $k \times |T|$ possible contracts with the colleges that were drawn in Step 1, where an acceptable permutation satisfies our assumption that students prefer higher levels of funding. Set the realized permutation as the student's preferences. Other contracts are not acceptable to the student.

For each realization of students' preferences, a (non-random) college admissions market is obtained. For notational convenience, we maintain the assumption of *uniform* random markets until the end of this section. In the appendix we show that our results continue to hold in a much broader class of preference distributions (Proposition 4). The natural generalization of the preference structure studied in Kojima and Pathak (2009), which allows for correlation in students' preferences, satisfies this condition, and so do cases where some fraction of the population considers certain levels of financial aid prohibitively low.[20]

A *sequence of uniform random markets*, denoted by $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$ where $\tilde{\Gamma}^n := \left\langle S^n, C^n, T^n, \{\succ_c, \gg_c, q_t^c\}_{t\in T^n}^{c\in C^n}, k^n \right\rangle$, is *regular* if there exist integers $k, l, \bar{q},$ and λ, all greater than one, such that:

1. $|C^n| = n$ for all n,
2. $k^n = k$ and $T^n = \{0, 1, ...l-1\}$ for all n,
3. $q_0^c \leq \bar{q}$ for all $c \in C^n$ and all n,
4. for all n, $c \in C^n$, and $s \in S^n$, $s \gg_c \emptyset$,
5. for all n and $c \in C^n$, there exist $t, t' \in T^n$ such that $q_t^c > q_{t'}^c > 0$, and
6. $\frac{1}{\lambda}n \leq |S^n| \leq \lambda n$, for all n.

Condition 1 assures that the number of colleges grows as the sequence progresses. Condition 2 assures that the number of contracts that students consider acceptable is uniformly bounded on the sequence. Condition 3 assures that the number of positions in each college is uniformly bounded across colleges and markets. Condition 4 assures that colleges find any student acceptable. These conditions are identical to those of Kojima and Pathak (2009).

Condition 5 is the key addition we make to their model. This condition assures, roughly, that each college faces a financial aid constraint on top of the capacity constraint in Condition 3. Finally, Condition 6 assures that the number of students does not grow much faster or much slower than the number of colleges. Kojima and Pathak (2009) only require the first half

[20] We also allow for some students who rank some pairs of contracts in a way that is not consistent with the natural order. This generalization captures situations where contracts are not naturally ranked (e.g., Aygün and Turhan, 2017), or when some applicants make mistakes (Hassidim, Romm and Shorrer, 2016; Artemov, Che and He, 2017; Shorrer and Sóvágó, 2017).

of this condition. Since we are interested in instances where a substantial fraction of colleges have multiple stable allocations, markets with a small number of students (who each find contracts with at most k schools acceptable) will clearly pose a problem as most schools will not even be a party to any individually rational allocation. Omitting the lower bound on $|S^n|$ and Condition 5 and requiring $|T| = 1$ would yield the definition of a sequence of uniform Kojima and Pathak markets.

3.2 Main theoretical results

Consider a regular sequence of uniform random markets, $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$. Let $\alpha(n)$ denote the expected number of students with multiple stable allocations, let $\beta(n)$ denote the expected number of colleges with multiple stable allocations that can successfully manipulate DA when all others are truthful, let $\gamma(n)$ denote the expected number of colleges with multiple stable allocations that cannot successfully manipulate DA when all others are truthful, let $\delta(n)$ denote the expected number of colleges with different numbers of assigned students in different core allocations, and let $\eta(n)$ denote the expected number of students who are matched in some stable allocation, but are unmatched in another; for simplicity, our notation suppresses the dependence on the sequence.

Theorem 2. *Given a regular sequence of uniform random markets, there exists $\Delta > 0$ such that:*

1. $\liminf_{n\to\infty} \alpha(n)/n > \Delta$,
2. $\liminf_{n\to\infty} \beta(n)/n > \Delta$,
3. $\liminf_{n\to\infty} \gamma(n)/n > \Delta$,
4. $\liminf_{n\to\infty} \delta(n)/n > \Delta$, *and*
5. $\liminf_{n\to\infty} \eta(n)/n > \Delta$.

To keep the focus on the crux of the argument we defer the detailed proof to the appendix and in what follows we analyze the special case of $T^n \equiv \{0, 1\}$ (e.g., students are either fully funded or not funded), $q_0^c = 2$ and $q_1^c = 1$ for all $c \in C^n$ for all n (i.e., each college has two seats and can offer one scholarship), and $|S^n| = 2|C^n|$ (i.e., there are as many students as there are college seats). We also defer the treatment of $\eta(n)$, which requires a more subtle argument. With this exception, the proof of the general case is heavier notationally, but is no more difficult conceptually than the argument we present below for the special case.

Consider two students, $r, p \in S^n$, a college $h \in C^n$, and some other college $c \in C^n$. Let the event $E^n(r, p, h, c)$ denote the case where:

1. College h ranks r higher than p on its master list. Formally, $r \gg_h p$.

2. The only students who find contracts with h acceptable are r and p. Formally, for all $s \in S^n \setminus \{r, p\}$ and all $t \in T^n$, $\emptyset \succ_s (s, h, t)$.

3. The only student who finds contracts with c acceptable is p. Formally, for all $s \in S^n \setminus \{p\}$ and all $t \in T^n$, $\emptyset \succ_s (s, c, t)$.

4. The two contracts r finds most desirable are $(r, h, 1)$ and $(r, h, 0)$, and r prefers the first to the second. Formally, for all $z \in X_r \setminus X_h$, $(h, r, 1) \succ_r (h, r, 0) \succ_r z$.

5. The two contracts p finds most desirable are $(p, h, 1)$ and $(p, c, 1)$, and p prefers the first to the second. Formally, for all $z \in X_p \setminus \{(p, h, 1), (p, c, 1)\}$, $(p, h, 1) \succ_p (p, c, 1) \succ_p z$.

Note that in the event $E^n(r, p, h, c)$, the students r and p and the colleges h and c have two stable allocations: $\{(r, h, 1), (p, c, 1)\}$ (which is their allocation in the outcome of DA with respect to the true preference profile), and $\{(r, h, 0), (p, h, 1)\}$. Also, note that college h can successfully manipulate DA by declaring that allocations under which r receives financial aid are not acceptable, but clearly c cannot. Furthermore, the two colleges have different numbers of students assigned to them in different core allocations.

Lemma 1. *Fix an arbitrary $\varepsilon > 0$. There exists sufficiently large n, such that for each college $h \in C^n$ the event*

$$E_h^n = \bigcup_{(r,p,c) \in S^n \times S^n \times C^n} E^n(r, p, h, c)$$

has probability bounded below by $\frac{\exp(-4k)}{k} - \varepsilon$.

Proof. For different selections of (r, p, c), the events $E^n(r, p, h, c)$ are disjoint. There are $2n \cdot (2n-1) \cdot (n-1)$ possible selections such that $r \neq p$ and $c \neq h$. Half of these events have zero probability (when $p \gg_h r$). The probability of each of the other events is greater than

$$(1 - \frac{k}{n-1})^{4n} \times \frac{1}{kn} \times \frac{1}{n}(1 - \frac{1}{k})\frac{1}{n},$$

where each term in this expression corresponds to an (independent)[21] requirement from the definition of $E^n(r, p, h, c)$. Thus, the probability of the event E_h^n is greater than $n \cdot (2n-1) \cdot (n-1) \times (1 - \frac{k}{n-1})^{4n} \times \frac{1}{kn} \times \frac{1}{n}(1 - \frac{1}{k})\frac{1}{n}$, and the limit of this expression is greater than or equal to $\frac{\exp(-4k)}{k}$.

□

The results on $\beta(n)$ and $\delta(n)$ follow immediately. The result on $\gamma(n)$ follows from a similar argument, where c is held fixed and the union is taken over selections of r, p, and h. The result on $\alpha(n)$ follows from the result on $\delta(n)$ and from the fact that each student finds contracts with at most k colleges acceptable.

4 Empirical Evidence

In this section we study the Hungarian college admissions market. We provide evidence in support of our assumptions on the demand structure (i.e., students' rank-ordered lists), corroborate the predictions of our model, and assess the potential welfare implications. Appendix C contains additional evidence from the Israeli Psychology Master's Match.

4.1 Background

College admissions in Hungary have been controlled centrally and organized through a centralized clearinghouse since 1985.[22] Each year, about 100,000 students apply to bachelor's programs and approximately 60,000 are assigned. As is standard in Europe, prospective students must choose a particular study program in advance (e.g., B.A. in applied economics at Corvinus University).

[21] To be precise, the first term refers to both the second and third requirements, as they are not independent.

[22] For more details see Biró (2007) and references therein.

Citizens of Hungary and of other member states of the European Economic Area are eligible to receive up to six years (12 semesters) of state-funded education. However, the government limits the number of state-funded seats in each field of study. While only eligible applicants may apply for admission with state funding, unfunded positions are also available and are open to all.[23]

Over the years, the mechanism used by the clearinghouse has changed several times. Since 2008, a variant of student-proposing DA has been in use. Prior to that, a variant of the contract-proposing algorithm was in place.[24] Both mechanisms endow applicants with field-specific priority scores based on a weighted average of several variables (mainly matriculation exam scores and GPA in the 11th and 12th grades, but also some credit for disabled, disadvantaged, or gifted applicants). The weights in the formula differ for different fields of study.

Since decisions about financial aid and admission are made simultaneously, and as the availability of funding may play a critical role in applicants' decisions between programs, applicants are allowed to submit rank-ordered lists (ROLs) ranking any number of contracts (program and funding-level combinations). For example, an applicant may submit an ROL that ranks three contracts with two programs: 1) a funded B.A. in applied economics at Corvinus University, 2) a funded B.Sc. in agricultural engineering at the University of Debrecen, and 3) an unfunded B.A. in applied economics at Corvinus University. Applicants who wish to submit an ROL that ranks more than three programs (corresponding to up to six contracts) are required to pay a fee (about $7 per additional program on the ROL). This feature implies that truthful reporting is not a dominant strategy in the Hungarian mechanism (Haeringer and Klijn, 2009). But given a set of programs on a list, it is dominated not to rank *all* acceptable contracts with these programs *truthfully*. To be clear, programs do not give precedence to applicants who apply for particular funding terms; they choose the highest-priority students available to them (under the restrictions on funding).

After the match results are realized, applicants are informed of their place-

[23]In 2013, tuition ranged from $2,000 to $23,000 for three years, with a median of $3,800 and a mean of $4,500. Many institutions grant funded students priority in access to subsidized housing and other amenities. The per capita GDP of Hungary in 2013 was $10,300.

[24]Under the contract-proposing version of DA, each contract is regarded as a separate program and the program-proposing variant of DA is used.

ment, and the *priority-score cutoff* for each contract is made public. The priority-score cutoff for a contract represents the lowest-score student who gained admission to the particular contract. These statistics receive extensive media coverage in the days after the match results are published.

4.2 Data

Our data on the Hungarian college admissions process between 2009 and 2011 is based on the dataset compiled by Shorrer and Sóvágó (2017), which merges data from four different sources. The main source of data is an administrative dataset containing much of the information available to the clearinghouse.[25] This dataset includes each applicant's ROL,[26] priority score in each relevant contract, as well as the information that is required to re-calculate it (i.e., academic performance in relevant exams and whether the applicant has a certified disadvantaged status).

The administrative data on students who applied during their senior year of high school is merged with the National Assessment of Basic Competencies dataset based on demographic variables. The NABC measures numeracy and literacy skills in a standardized way. Since 2008 it has covered all students in the 6th, 8th, and 10th grades who attended school on the day of the exam (prior to 2008 it covered a sample). The NABC dataset includes self-reported survey measures of socioeconomic status (e.g., parental education, home possessions, etc.). Shorrer and Sóvágó (2017) construct an NABC-based socioeconomic status (SES) index following Horn (2013). This index consists of three sub-indices: an index of parental education, an index of home possessions (number of bedrooms, cars, books, computers, etc.), and an index of parental labor market status.[27]

Shorrer and Sóvágó's dataset also includes microregion-level annual unemployment rates published by the National Employment Service in 2008, with a territorial breakdown consisting of 174 units. Finally, it includes the

[25]The Hungarian Higher Education Application Database (FELVI) is owned by the Hungarian Education Bureau (Oktatasi Hivatal). The data were processed by the Hungarian Academy of Sciences Centre for Economic and Regional Studies (HAS-CERS).

[26]The dataset reports the first 6 contracts on an applicant's ROL as well as the applicant's allocation, in case it was ranked lower. It also specifies the number of contracts on each ROL. In 93.6% of the cases, we observe the complete ROL.

[27]The objectives of the NABC are similar to those of the OECD Program for International Student Assessment (PISA). The NABC-based socioeconomic status index resembles the PISA economic, social, and cultural status (ESCS) indicator.

per capita gross annual income for all 3,164 settlements for each year of the sample, calculated based on information published by the Hungarian Central Statistics Office.

4.3 Student rank-ordered lists

We now ask: *Is the data consistent with the assumptions of our model on the distribution of student ROLs?* Namely, are ROLs "short" (Roth and Peranson, 1999), and are applicants from lower socioeconomic backgrounds more likely to rank funded positions only, as our leading example suggests? The answer is positive.

We find that 93.6% of the ROLs rank up to six contracts, and 99.1% of all ROLs are shorter than 10 contracts. In Table 1 we compare the characteristics of applicants who submitted ROLs ranking funded contracts exclusively, to those of applicants who submitted ROLs ranking both funded and unfunded contracts.[28] We find that, on average, applicants who ranked funded contracts exclusively come from lower SES backgrounds, and have lower academic ability.

The fact that applicants with stronger academic performance were more likely to submit an ROL ranking funded contracts exclusively can be explained by these applicants expecting to be able to gain admission to a funded contract. In case they are not admitted with funding, they may prefer to retake some exams and reapply the following year rather than pay tuition (Krishna, Lychagin and Robles, 2015), or they may be optimistic enough about their chances of admission to be nearly indifferent between their ROL and another ROL that ranks unfunded contracts (Chen and Pereyra Barreiro, 2015; Artemov, Che and He, 2017). Since SES is positively correlated with academic ability (Table 2), this pattern likely leads our analysis to understate the true scope of the gains from reallocating financial aid from wealthy applicants to needy ones.

[28] The third group, applicants who submitted ROLs ranking only unfunded contracts, is small and includes ineligible students (e.g., applicants from countries outside the European Economic Area) and applicants who made mistakes (Rees-Jones, 2015; Hassidim et al., 2017).

Table 1: Characteristics of applicants who submitted ROLs with funded contracts only

Dependent variable	Funded contracts only						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NABC-based SES index	-0.062***	-0.069***					
	(0.0017)	(0.0017)					
11th-grade GPA (1-5)		0.077***		0.094***		0.094***	0.093***
		(0.0023)		(0.0012)		(0.0012)	(0.0012)
Income (1000 USD)			-0.038***	-0.039***			
			(0.0006)	(0.0006)			
Unemployment (%)					0.008***	0.008***	
					(0.0002)	(0.0002)	
Observations	78064	78064	284701	284701	284701	284701	284701
R^2	0.017	0.032	0.016	0.038	0.007	0.028	0.021

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The regression coefficients are conditional on a year fixed effect and an indicator for missing values. Robust standard errors are in parentheses. The sample includes all ROLs excluding those that ranked unfunded contracts only. In Columns 1 and 2 we restrict the sample to high-school-senior applicants, the population that was matched to the NABC data. The NABC-based SES index was matched to the main dataset based on 5 variables (year and month of birth, gender, school identifier, and four-digit postal code). It is normalized to have a mean of 0 and a standard deviation of 1 in the population of high-school students. Income stands for the per capita gross annual income in the settlement where the applicant resided.

Table 2: Academic ability and socioeconomic status

Dependent variable	Grade-11 GPA		
	(1)	(2)	(3)
NABC-based SES index	0.085***		
	(0.0028)		
Unemployment (%)		-0.005***	
		(0.0004)	
Income (1000 USD)			0.016***
			(0.0011)
Observations	78133	284701	284701
R^2	0.325	0.056	0.056

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The regression coefficients are conditional on a year fixed effect and an indicator for missing values. Robust standard errors are in parentheses. The sample includes all ROLs excluding those that ranked unfunded contracts only. In Column 1 we restrict the sample to high-school-senior applicants, the population that was matched to the NABC data. The NABC-based SES index was matched to the main dataset based on 5 variables (year and month of birth, gender, school identifier, and four-digit postal code). It is normalized to have a mean of 0 and a standard deviation of 1 in the population of high-school students. Income stands for the per capita gross annual income in the settlement where the applicant resided.

4.4 The core

We proceed to the main question of interest: *Is the core of the Hungarian college admissions market large, and can colleges successfully manipulate the mechanism?* As we have established above, the core of a college admissions market does not, generally, admit a lattice structure with respect to same-side preferences. Thus, we are unable to characterize it fully using standard methods. Moreover, data limitations further complicate our ability to verify stability.

Instead, we assume that the strategic unit on the colleges side is a field of study – the unit that shares priorities and budget – and that its priorities, which are determined by the government, represent the field's true preferences. We ask how much each field can improve its yield by applying market power over students who are placed in the field under DA when all colleges are truthful. Since each field of study offers multiple programs in multiple locations, we conservatively assume only that each field is indifferent between funding levels given a student's placement, so long as the field's budget is not exceeded. Specifically, we do not take a stance on the field's preferences with respect to transferring students from one program to another program in the field.

Given a field, f, and a year, t, we focus on identifying students who receive financial aid in f even though the next-highest-ranked contract on their ROL whose priority score cutoff they pass is the unfunded contract with the same program (i.e., the same study track in the same institution). The collection of such students, MP_f^t, is a set of students over which f has market power: f can safely refuse their funding.

There are 10,056 students in some MP_f^t in our sample period. They correspond to approximately 8% of the tuition waivers offered by the state in this period. Namely, holding the financial aid offered in all other programs constant, about 8% of the waivers have no effect on the receiving students' choice of program.

We are interested in knowing if by refusing funding to students in MP_f^t the field can improve the incoming cohort. We are especially interested in knowing if such behavior will increase the total number of students attending *some* college. To this end, we define the set DB_f^t.

Definition 1. Given a program, f, and a year, t, let DB_f^t be the collection of up to $|MP_f^t|$ highest priority score year-t applicants who were unassigned or assigned to a contract (not with f) that they ranked lower than the funded contract with a program in f that had free capacity.

Since we do not have data on exact capacities, we say that a program has free capacity if it did not reject any applicant to an unfunded position. We take a partial-equilibrium approach and assume that if f has free capacity in some program, it can apply market power over members of MP_f^t and use the freed-up funds to admit students in DB_f^t, and the resulting allocation will be stable.[29] Students in DB_f^t stand to *directly benefit* if f applies market power.

This approach is conservative in several ways. First, we do not consider improvements that the field can make by coordinating across its various programs. Second, we ignore programs in which the field can improve only the composition of the student body without affecting the size of the incoming cohort. Had each program (or college) acted as an independent strategic unit, this second source of gains would have been substantial, as a substantial fraction of applicants apply only to programs in one field.

We perform the analysis separately for each field in each year. We find that 9,463 students stand to benefit directly. Of those, 5,886 students were not placed in any college in practice (i.e., under DA). Table 3 compares the characteristics of the groups. We find that the students who would benefit from moving to another stable allocation (i.e., members of some DB_f^t), on average, come from lower SES relative to those who stand to lose from such a change (i.e., members of some MP_f^t), that they are more likely to live in a village, and that they are less likely to live in the capital, Budapest. They are also more likely to be female and to have graduated from a vocational high school. Mechanically, winners also have lower academic achievements.

5 Discussion

We have established that a centralized two-sided college admissions market that uses the student-proposing version of DA leaves much room for colleges to strategize. This finding stands in sharp contrast to the findings in markets without contracts. Our results suggest that colleges have an incentive to

[29]The rationale behind this approach derives from Kojima and Pathak (2009). General equilibrium effects on f would require an unlikely "cycle of vacancies" given the demand structure.

Table 3: Characteristics of applicants in MP and DB

	DB		MP	
	mean	sd	mean	sd
Disadvantaged (dummy)	0.09	0.288	0.03	0.173
Unemployment rate (%)	7.95	4.668	6.77	4.162
Gross annual per capita income (1000 USD)	6.03	1.451	6.59	1.540
11th-grade GPA	3.77	0.777	3.90	0.793
Female	0.58	0.493	0.50	0.500
Secondary grammar school	0.64	0.479	0.68	0.467
Vocational school	0.32	0.468	0.26	0.439
Capital	0.14	0.348	0.26	0.439
County capital	0.21	0.405	0.21	0.406
Town	0.34	0.474	0.29	0.452
Village	0.31	0.463	0.24	0.430
Programs in ROL	3.31	1.295	2.25	1.009
Contracts in ROL	3.70	1.773	4.21	2.273
Observations	9,463		10,056	
Unassigned under DA	5,886		0	

Notes: The table compares the characteristics of applicants over whom some field can apply market power (members of some MP_f^t) to those of students who stand to benefit directly from moving to another stable matching allocation (members of some DB_f^t). The sample covers the years between 2009 and 2011. Each year, approximately 60,000 students are assigned to college through DA, of which approximately 70% are funded.

provide financial aid based on need rather than merit, even if they do not have preferences for equity or social justice and their only goal is simply to maximize the quality of their incoming cohort.

We have also shown that using student-proposing DA in a college admissions market is akin to allocating financial aid based on merit. This result generalizes to contract-proposing DA, where the algorithm regards each contract as a separate college – the mechanism used in Turkish college admissions (Balinski and Sönmez, 1999).[30] A report by the World Bank states that the situation in this country "is akin to giving a large number of scholarships in each institution on the basis of merit." The report adds that "merit-based scholarships to be funded by government make sense if high calibre students need extra inducements for entering higher education. It is not obvious that Turkish students need such inducement," and that "an important group to target would be students from less privileged backgrounds, either in terms of income, regions, ethnicity or gender." Our findings provide support for these assertions and offer guidance on how to implement alternative policies.

Finally, we have shown that different stable allocations in college admissions markets differ substantially. Thus, there are economically meaningful trade-offs even when concentrating on the core. This observation suggests that answering "classic" questions in the theory of two-sided matching markets for the college admissions setting may be a fruitful direction. A natural question, for example, is: *How to find the stable matching that matches the most students?* Understanding the structure of the core in college admissions environments is, thus, a promising research direction.

References

Abizada, Azar. 2016. "Stability and incentives for college admissions with budget constraints." *Theoretical Economics*, 11(2): 735–756.

Afacan, Mustafa Oguz. 2017. "Graduate Admission with Financial Support." mimeo.

[30]In Turkey, private colleges are obligated by law to offer a full scholarship to, at least, 15% of their student body. Additionally, most institutions offer multiple contracts in the same program: a subsidized morning schedule, and a more expensive evening schedule (Akar, 2010; Hatakenaka, 2006).

Akar, Hanife. 2010. “Globalization and its challenges for developing countries: the case of Turkish higher education.” *Asia Pacific Education Review*, 11(3): 447–457. Copyright - Education Research Institute, Seoul National University, Seoul, Korea 2010; Last updated - 2015-12-19.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2017. “Strategic Mistakes: Implications for Market Design Research.”

Ashlagi, Itai, Yash Kanoria, and Jacob D Leshno. 2017. “Unbalanced random matching markets: The stark effect of competition.” *Journal of Political Economy*, 125(1): 69–98.

Avery, Christopher, and Caroline Minter Hoxby. 2004. “Do and should financial aid packages affect students’ college choices?” In *College choices: The economics of where to go, when to go, and how to pay for it*. 239–302. University of Chicago Press.

Aygün, Orhan, and Bertan Turhan. 2017. “Large Scale Affirmative Action in School Choice: Admissions to IITs in India.” mimeo.

Aygün, Orhan, and Inácio Bó. 2016. “College admission with multidimensional privileges: The Brazilian affirmative action case.” *Unpublished paper, WZB Berlin Social Science Center.[869].*

Aygün, Orhan, and Tayfun Sönmez. 2013. “Matching with contracts: Comment.” *American Economic Review*, 103(5): 2050–2051.

Azevedo, Eduardo M., and Eric Budish. 2012. “Strategyproofness in the large as a desideratum for market design.” 55–55, ACM.

Azevedo, Eduardo M., and Jacob D. Leshno. 2016. “A supply and demand framework for two-sided matching markets.” *Journal of Political Economy*, 124(5): 1235–1268.

Balinski, M., and T. Sönmez. 1999. “A tale of two mechanisms: student placement.” *Journal of Economic Theory*, 84(1): 73–94.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune. 2013. “Marry for what? Caste and mate selection in modern India.” *American Economic Journal: Microeconomics*, 5(2): 33–72.

Biró, Péter. 2007. "Higher education admission in Hungary by a score-limit algorithm."

Chen, Li, and Juan Sebastian Pereyra Barreiro. 2015. "Self-Selection in School Choice." ECARES Working Papers.

Coles, Peter, and Ran Shorrer. 2014. "Optimal truncation in matching markets." *Games and Economic Behavior*, 87: 591–615.

Coles, Peter, Yannai Gonczarowski, and Ran Shorrer. 2014. "Strategic Behavior in Unbalanced Matching Markets." mimeo.

Demange, Gabrielle, David Gale, and Marilda Sotomayor. 1986. "Multi-item auctions." *The Journal of Political Economy*, 863–872.

DePalma, Anthony. 1992. "M.I.T. ruled guilty in antitrust case." *The New York Times*.

Dimakopoulos, Philipp D., and C.-Philipp Heller. 2015. "Matching with waiting times: The german entry-level labour market for lawyers." Working paper.

Dubins, Lester E., and David A. Freedman. 1981. "Machiavelli and the Gale–Shapley algorithm." *American Mathematical Monthly*, 88(7): 485–494.

Dur, Umut M, Scott Duke Kominers, Parag A Pathak, and Tayfun Sönmez. 2013. "The demise of walk zones in Boston: Priorities vs. precedence in school choice." National Bureau of Economic Research.

Dur, Umut, Parag A Pathak, and Tayfun Sönmez. 2016. "Explicit vs. Statistical Preferential Treatment in Affirmative Action: Theory and Evidence from Chicago's Exam Schools." National Bureau of Economic Research.

Echenique, Federico. 2012. "Contracts versus salaries in matching." *American Economic Review*, 102(1): 594–601.

Ehlers, Lars. 2004. "In search of advice for participants in matching markets which use the deferred-acceptance algorithm." *Games and Economic Behavior*, 48(2): 249 – 270.

Ehlers, Lars, and Jordi Massó. 2007. "Incomplete information and singleton cores in matching markets." *Journal of Economic Theory*, 136(1): 587–600.

Fleiner, Tamás. 2003. "A Fixed-Point Approach to Stable Matchings and Some Applications." *Mathematics of Operations Research*, 28(1): 103–126.

Gale, David, and Lloyd S. Shapley. 1962. "College admissions and the stability of marriage." *American Mathematical Monthly*, 69(1): 9–15.

Haeringer, Guillaume, and Flip Klijn. 2009. "Constrained school choice." *Journal of Economic Theory*, 144(5): 1921–1947.

Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer. 2016. "'Strategic' Behavior in a Strategy-proof Environment." Working paper.

Hassidim, Avinatan, Assaf Romm, and Ran I Shorrer. 2017. "Redesigning the Israeli Psychology Master's Match." *American Economic Review*, 107(5): 205–209.

Hassidim, Avinatan, Marciano Déborah, Assaf Romm, and Ran I. Shorrer. 2017. "The mechanism is truthful, why aren't you?" *American Economic Review: Papers and Proceedings*, 107(5).

Hatakenaka, Sachi. 2006. "Higher education in Turkey for 21st century: Size and composition." *World Bank. Retrieved February*, 25(2009): 444607–1192636551820.

Hatfield, John W., and Fuhito Kojima. 2010. "Substitutes and stability for matching with contracts." *Journal of Economic Theory*, 145(5): 1704–1723.

Hatfield, John W., and Paul R. Milgrom. 2005. "Matching with contracts." *American Economic Review*, 95(4): 913–935.

Hatfield, John W., and Scott D. Kominers. 2015. "Hidden Substitutes." Proceedings of the 16th ACM Conference on Economics and Computation, 37.

Hatfield, John W., Scott D. Kominers, and Alexander Westkamp. 2015. "Stability, strategy-proofness, and cumulative offer mechanisms." mimeo.

Hitsch, Günter, Ali Hortaçsu, and Dan Ariely. 2010. "Matching and Sorting in Online Dating." *The American Economic Review*, 100(1): 130–163.

Holzman, Ron, and Dov Samet. 2014. "Matching of like rank and the size of the core in the marriage problem." *Games and Economic Behavior*, 88: 277–285.

Horn, Dániel. 2013. "Diverging performances: the detrimental effects of early educational selection on equality of opportunity in Hungary." *Research in Social Stratification and Mobility*, 32: 25–43.

Hoxby, Caroline M. 2000. "Benevolent colluders? The effects of antitrust action on college financial aid and tuition." National bureau of economic research.

Immorlica, Nicole, and Mohammad Mahdian. 2005. "Marriage, honesty, and stability." 53–62, Society for Industrial and Applied Mathematics.

Jagadeesan, Ravi. 2016. "Cadet-branch matching in a quasi-linear labor market." 61–62, ACM.

Kelso, Jr., Alexander S., and Vincent P. Crawford. 1982. "Job Matching, Coalition Formation, and Gross Substitutes." *Econometrica*, 50(6): 1483–1504.

Klaus, Bettina, and Flip Klijn. 2006. "Median stable matching for college admissions." *International Journal of Game Theory*, 34(1): 1–11.

Knuth, Donald E. 1976. *Mariages stables et leurs relations avec d&autres problèmes combinatoires.* Presses de l'Université de Montréal.

Kojima, Fuhito, and Parag A. Pathak. 2009. "Incentives and Stability in Large Two-Sided Matching Markets." *American Economic Review*, 99(3): 608–27.

Kominers, Scott Duke. 2012. "On the correspondence of contracts to salaries in (many-to-many) matching." *Games and Economic Behavior*, 75(2): 984 – 989.

Konishi, Hideo, and M. Ünver. 2006. "Games of Capacity Manipulation in Hospital-intern Markets." *Social Choice & Welfare*, 27(1): 3 – 24.

Krishna, Kala, Sergey Lychagin, and Verónica Frisancho Robles. 2015. "Retaking in High Stakes Exams: Is Less More?" National Bureau of Economic Research.

Lee, SangMok. 2017. "Incentive compatibility of large centralized matching markets." *The Review of Economic Studies*.

Menzel, Konrad. 2015. "Large Matching Markets as Two-Sided Demand Systems." *Econometrica*, 83(3): 897–941.

Niederle, Muriel. 2007. "Competitive Wages in a Match with Ordered Contracts." *American Economic Review*, 97(5): 1957–1969.

Pakzad-Hurson, Bobak. 2014. "Stable and efficient resource allocation with contracts." Working Paper.

Pathak, Parag A, and Tayfun Sönmez. 2008. "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism." *American Economic Review*, 98(4): 1636–1652.

Pittel, Boris. 1989. "The average number of stable matchings." *SIAM Journal on Discrete Mathematics*, 2(4): 530–549.

Pittel, Boris. 1992. "On likely solutions of a stable marriage problem." *The Annals of Applied Probability*, 2(2): 358–401.

Rees-Jones, Alex. 2015. "Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match." Available at SSRN 2662670.

Roth, Alvin E. 1982. "The economics of matching: Stability and incentives." *Mathematics of Operations Research*, 7(4): 617–628.

Roth, Alvin E. 1984*a*. "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory." *Journal of Political Economy*, 92(6): 991–1016.

Roth, Alvin E. 1984*b*. "Stability and polarization of interests in job matching." *Econometrica: Journal of the Econometric Society*, 47–57.

Roth, Alvin E. 1986. "On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Markets." *Econometrica*, 54(2): 425–427.

Roth, Alvin E. 2002. "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica*, 70(4): 1341–1378.

Roth, Alvin E, and Elliott Peranson. 1999. "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic D." *American economic review*, 89(4): 748–782.

Roth, Alvin E., and Uriel G. Rothblum. 1999. "Truncation Strategies in Matching Markets-in Search of Advice for Participants." *Econometrica*, 67(1): 21–43.

Roth, Alvin E, and Xiaolin Xing. 1994. "Jumping the gun: Imperfections and institutions related to the timing of market transactions." *The American Economic Review*, 992–1044.

Schlegel, Jan Christoph. 2015. "Contracts versus salaries in matching: A general result." *Journal of Economic Theory*, 159: 552–573.

Schwarz, Michael, and M Bumin Yenmez. 2011. "Median stable matching for markets with wages." *Journal of Economic Theory*, 146(2): 619–637.

Shorrer, Ran I., and Sándor Sóvágó. 2017. "Obvious Mistakes in a Strategically Simple College-Admissions Environment." mimeo.

Sönmez, Tayfun. 1997. "Manipulation via Capacities in Two-Sided Matching Markets." *Journal of Economic Theory*, 77(1): 197 – 204.

Sönmez, Tayfun. 1999. "Strategy-Proofness and Essentially Single-Valued Cores." *Econometrica*, 67(3): 677–689.

Sönmez, Tayfun. 2013. "Bidding for army career specialties: Improving the ROTC branching mechanism." *Journal of Political Economy*, 121(1): 186–219.

Sönmez, Tayfun, and Tobias B. Switzer. 2013. "Matching with (branch-of-choice) contracts at the United States Military Academy." *Econometrica*, 81(2): 451–488.

Storms, Evan. 2013. "Incentives and Manipulation in Large Market Matching with Substitutes." Mimeo.

Teo, Chung-Piaw, and Jay Sethuraman. 1998. "The geometry of fractional stable matchings and its applications." *Mathematics of Operations Research*, 23(4): 874–891.

Westkamp, Alexander. 2013. "An analysis of the German university admissions system." *Economic Theory*, 53(3): 561–589.

Yenmez, M Bumin. 2015. "College admissions."

A Additional Examples

We extend Example 3 to show that a college may possess market power over a student, even if the student does not rank the contracts with the college contiguously.

Example 4. We add to Example 3 another (elite) college e, and a (genius) student g. The elite college has one seat and one level of funding ($q_0^e = 1$, $q_1^e = 0$), and its preferences are summarized by

$$(g, e, 0) \succ_e (r, e, 0) \succ_e (p, e, 0) \succ_e \emptyset.$$

Student g's first-choice college is e. Her preferences are given by

$$(g, e, 0) \succ_g (g, h, 1) \succ_g (g, h, 0) \succ_g (g, c, 0) \succ_g \emptyset.$$

Other students' preferences are now

$$(r, h, 1) \succ_r (r, e, 0) \succ_r (r, h, 0) \succ_r (r, c, 1) \succ_r (r, c, 0) \succ_r \emptyset,$$

and

$$(p, h, 1) \succ_p (p, c, 1) \succ_p (p, c, 0) \succ_p (p, h, 0) \succ_p (p, e, 0) \succ_p \emptyset.$$

Thus, the elite college and the genius applicant are the first choice of one another, and hence must be matched in any stable allocation. There are two stable allocations under these preferences: $\{(r, h, 1), (p, c, 0), (g, e, 0)\}$, which is the result of the student-proposing DA algorithm, and $\{(r, h, 0), (p, h, 1), (g, e, 0)\}$.

Had the elite college been interested in the rich applicant (e.g., if its capacity was 2), then the unique core allocation would be $\{(r, h, 1), (p, c, 0), (g, e, 0)\}$. Intuitively, the fact that the rich applicant prefers the elite institution to the unfunded position at h, combined with the institution's willingness to accept the rich applicant, eliminates h's ability to apply market power over the rich applicant. That h can apply market power over r is a result of e not being interested in r, which means that r does not have an outside option he prefers.

B Proofs

We now prove the theorem for the general case. We require some additional notation.

Given a college c in C^n, let $\bar{t}_c$ be the maximal $t \in T^n$ such that $0 < q^c_{\bar{t}_c} < q^c_{\bar{t}_c - 1}$. Similarly, let $\hat{t}_c$ be the maximal $t \in T^n$ such that $0 < q^c_{\hat{t}_c}$. The existence of $\bar{t}_c$ and $\hat{t}_c$ is assured by Condition 5 in the definition of a sequence of uniform random markets.

Given a college h in C^n, an ordered selection of $\bar{t}_h + 1$ students in S^n, $(p, r_1, r_2, ...r_{q^h_{\bar{t}_h}})$, and another college, c in C^n, let the event $E^n(p, r_1, r_2, ...r_{q^h_{\bar{t}_h}}, h, c)$ denote the case where:

1. College h ranks lower-index r_js higher on its master list, and ranks all r_js higher than p. Formally, $r_j \gg_h r_i \gg_h p$ for all $1 \leq j < i \leq q^h_{\bar{t}_h}$.

2. The only students who find contracts with h acceptable are the members of $\{p, r_1, r_2, ...r_{q^h_{\bar{t}_h}}\}$. Formally, for all $s \in S^n \setminus \{p, r_1, r_2, ...r_{q^h_{\bar{t}_h}}\}$ and all $t \in T^n$, $\emptyset \succ_s (s, h, t)$.

3. The only student who finds contracts with c acceptable is p. Formally, for all $s \in S^n \setminus \{p\}$ and all $t \in T^n$, $\emptyset \succ_s (s, c, t)$.

4. All r_j's prefer to be placed in h under any terms to any contract with another college. Formally, for all $1 \leq j \leq q^h_{\bar{t}_h}$, for all $z \in X_{r_j} \setminus X_h$, $(h, r_j, 0) \succ_{r_j} z$.

5. The most desirable contracts for p are $(p, h, |T|-1)$,$(p, h, |T|-2)$,...,$(p, h, \bar{t}_h)$, followed by $(p, c, |T|-1)$,$(p, c, |T|-2)$,...,$(p, c, 0)$. Formally, for all $z \in X_p$, $z \succ_p (p, c, |T|-1)$ if and only if $z \in \{(p, h, t) | t \geq \bar{t}_h\}$, and $z \succ_p (p, c, 0)$ if and only if $z \in \{(p, h, t) | t \geq \bar{t}_h\} \cup \{(p, c, t) | t > 0\}$.

Note that in the event $E^n(p, r_1, r_2, ...r_{q^h_{\bar{t}_h}}, h, c)$, student p and colleges h and c have multiple stable allocations. The stable allocation resulting from

DA includes $(r_1, h, \hat{t}_h)$ and $(p, c, \hat{t}_c)$. But another stable allocation involves a contract of the form (p, h, t) for $t \geq \bar{t}_h$ (and some r_j receiving a lower level of financial aid at h). Note that college h can successfully manipulate DA (e.g., by declaring that allocations under which r_1 receives financial aid are not acceptable), but clearly c cannot. Furthermore, the two colleges have different numbers of students assigned to them in different core allocations.

Lemma 2. *There exists $L > 0$ and n', such that for all $n > n'$ and for each college $h \in C^n$, the event*

$$E_h^n = \bigcup_{(p, r_1, r_2, \ldots r_{q^h_{\bar{t}_h}}, c) \in S^n \times \ldots \times S^n \times C^n} E^n(p, r_1, r_2, \ldots r_{q^h_{\bar{t}_h}}, h, c)$$

has probability bounded below by L.

Proof. For sufficiently large n, given a selection of h, there are at least $\binom{\lfloor \frac{n}{\lambda} \rfloor}{q^h_{\bar{t}_h}+1}$ selections of $(p, r_1, r_2, \ldots r_{q^h_{\bar{t}_h}})$ that meet the first condition. Given such a selection and a selection of one of the $n-1$ other colleges $c \neq h$, the probability that the other, independent[31] conditions are satisfied is bounded below by

$$(1 - \frac{k}{n-1})^{\lceil 2\lambda n \rceil} \times \frac{1}{n \cdot k^{|T|-1}} \times \frac{1}{n^2 \cdot k^{2|T|}}.$$

Moreover, since, given h, the events $E^n(p, r_1, r_2, \ldots r_{q^h_{\bar{t}_h}}, h, c)$ are disjoint, the probability of their (disjoint) union is greater than

$$(n-1)\binom{\lfloor \frac{n}{\lambda} \rfloor}{q^h_{\bar{t}_h}+1} \times (1 - \frac{k}{n-1})^{\lceil 2\lambda n \rceil} \times \frac{1}{(n \cdot k^{|T|-1})^{q^h_{\bar{t}_h}}} \times \frac{1}{n^2 \cdot k^{2|T|}},$$

which converges to a positive constant (that depends on $q^h_{\bar{t}_h}$) as n grows large. Since $q^h_{\bar{t}_h} < \bar{q}$ by Condition 3 in the definition of uniform random markets, $q^h_{\bar{t}_h}$ can take only finitely many values; hence taking the minimum of the limits and subtracting some small ε suffices.

□

Proof (of Theorem 2). The results on $\beta(n)$ and $\delta(n)$ follow directly from the lemma. Let $\chi_{E_h^n}$ denote the indicator of the event $\chi_{E_h^n}$. The results on

[31] To be precise, the second and third conditions are not independent. The first term of our bound applies to both of them simultaneously.

$\gamma(n)$ follow by noting that the expectation of the random variable $\sum_{h\in C^n} \chi_{E_h^n}$ increases linearly in n, which implies that the expected number of colleges "playing the role of c" is large. The result on $\alpha(n)$ follows from the result on $\delta(n)$, and from the fact that each student has at most k colleges she prefers to the outside option. The result on $\eta(n)$ requires more work, and follows from Lemma 4 below. □

Lemma 3. *Given a profile of (complete) college master lists, at least $\frac{1}{3}$ of the students are ranked between $\frac{1}{4}|S|$ and $\frac{3}{4}|S|$ in at least $\frac{1}{4}$ of the lists.*

Proof. The fraction of students who appear in this half of the list of at least $\frac{1}{4}$ of the lists is equal to the probability that this condition is satisfied by a student drawn uniformly at random. Let $\chi_c(\cdot)$ denote the indicator variable that a student is in the top or bottom quarter of c's list. Then, by Markov's inequality,

$$\Pr\{\sum_{c\in C}\chi_c > \frac{3}{4}n\} \leq \frac{n/2}{3n/4} = \frac{4}{6},$$

which completes the proof. □

Denote by $M^n \subset S^n$ the collection of students who are ranked between $\frac{1}{4}|S^n|$ and $\frac{3}{4}|S^n|$ in at least $\frac{n}{4}$ of the lists. Consider a student in M^n and k colleges on whose master list the student is ranked between $\frac{1}{4}|S^n|$ and $\frac{3}{4}|S^n|$. Denote one of the colleges by $h \in C^n$, the student by $r_{q^h_{\bar{t}_h}} \in S^n$, and the other colleges by $\{c_i\}_{i=1}^{k-1} \in C^n$.

Select an additional college $c \in C^n$ and q_0^h students in S^n, $(p, r_1, r_2, ..., r_{q^h_{\bar{t}_h}-1}, r_{q^h_{\bar{t}_h}+1}, ..., r_{q_0^h})$. Finally, for each c_i select $\bar{q}$ different students whom the college ranks among the highest $\frac{1}{4}|S^n|$, $\{s^j_{c_i}\}_{j=1}^{\bar{q}}$. Let the event $E^n\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s^j_{c_i}\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ denote the case where:

1. College h ranks lower-index r_js higher on its master list, and ranks p between $r_{q^h_{\bar{t}_h}}$ and $r_{q^h_{\bar{t}_h}+1}$. Formally, $r_j \gg_h r_i$ for all $1 \leq j < i \leq q^h_{\bar{t}_h}$ and $r_{q^h_{\bar{t}_h}} \gg_h p \gg_h r_{q^h_{\bar{t}_h}+1}$.

2. The only students who find contracts with h acceptable are the members of $\{p, r_1, r_2, ...r_{q_0^h}\}$. Formally, for all $s \in S^n \setminus \{p, r_1, r_2, ...r_{q_0^h}\}$ and all $t \in T^n$, $\emptyset \succ_s (s, h, t)$.

3. The only student who finds contracts with c acceptable is p. Formally, for all $s \in S^n \setminus \{p\}$ and for all $t \in T^n$, $\emptyset \succ_s (s, c, t)$.

4. All r_js prefer to be placed in h under any terms to any contract with another college. Formally, for all $1 \le j \le q_0^h$, for all $z \in X_{r_j} \setminus X_h$, $(h, r_j, 0) \succ_{r_j} z$.

5. The most desirable contracts for p are $(p, h, |T|-1)$,$(p, h, |T|-2)$,...,$(p, h, \bar{t}_h)$, followed by $(p, c, |T|-1)$,$(p, c, |T|-2)$,...,$(p, c, 0)$. Formally, for all $z \in X_p \setminus \{(p, h, t)|t \ge \bar{t}_h\}$, $(p, c, |T|-1) \succ_p z$, and $z \succ_p (p, c, 0)$ if and only if $z \in \{(p, h, t)|t \ge \bar{t}_h\} \cup \{(p, c, t)|t > 0\}$.

6. The lowest-ranked r-student, $r_{q_0^h}$, finds contracts only with $(h, c_1, c_2, ...c_{k-1})$ acceptable, and prefers $(r_{q_0^h}, c_i, |T|-1)$ to $(r_{q_0^h}, c_j, |T|-1)$ if $i < j$. Formally, for all $u \in C^n$ and $t \in T^n$, $u \notin \{h, c_1, c_2, ...c_{k-1}\} \implies \emptyset \succ_{r_{q_0^h}} (r_{q_0^h}, u, t)$, and $(r, c_i, |T|-1) \succ_{r_{q_0^h}} (r, c_j, |T|-1)$ if $i < j$ if and only if $i < j$.

7. For each c_i the only students who find contracts with c_i acceptable are the members of $\{s_{c_i}^j\}_{j=1}^{\bar{q}}$. Formally, for all $i \in \{1, 2, ..., k-1\}$, for all $s \in S^n \setminus \{s_{c_i}^j\}_{j=1}^{\bar{q}}$, and for all $t \in T^n$, $\emptyset \succ_s (s, c_i, t)$.

8. For each c_i the members of $\{s_{c_i}^j\}_{j=1}^{\bar{q}}$ prefer to be placed in c_i under any terms to any contract with another college. Formally, for all $i \in \{1, 2, ..., k-1\}$, for all $1 \le j \le \bar{q}$, for all $z \in X_{s_{c_i}^j} \setminus X_{c_i}$, $(c, s_{c_i}^j, 0) \succ_{s_{c_i}^j} z$.

Note that in the event $E^n\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ the stable outcome that corresponds to DA assigns students $\{r_1, r_2, ...r_{q_0^h}\}$ to h, and student p to c. But in another stable allocation, college c and student $r_{q_0^h}$ receive no assignment, and students $\{p, r_1, r_2, ...r_{q_0^h-1}\}$ are assigned to h, where one of the students in $\{r_1, r_2, ...r_{q_{\bar{t}_h}^h}\}$ receives a lower level of financial aid relative to the stable outcome that corresponds to DA.

Lemma 4. *There exists $L > 0$ and n', such that for $n > n'$, the probability that the event $E^n\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ occurs for some valid[32] selection of arguments is greater than L.*

[32] We use the term "valid" when we refer to conditions on college preferences, to emphasize that these are arbitrary (i.e., not assumed random).

Proof. Conditional on the selection of $\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ being valid, the probability of the event is bounded below by

$$(1-\frac{k}{n-2k})^{2\lambda n}\times(\frac{1}{k})^{|T|\bar{q}}\cdot(\frac{1}{n})^{q_0^h}\times(\frac{1}{k})^{2|T|}\cdot(\frac{1}{n})^{2}\times(\frac{1}{n})^{k-1}\times(1-\frac{k}{n-2k})^{k\lambda n}\cdot(\frac{1}{n})^{\bar{q}(k-1)}\times(\frac{1}{k})^{k|T|}.$$

This expression behaves asymptotically like

$$\bar{C}\cdot(\frac{1}{n})^{q_0^h}\cdot(\frac{1}{n})^{k+1}\cdot(\frac{1}{n})^{\bar{q}(k-1)} = \bar{C}\cdot(\frac{1}{n})^{q_0^h+\bar{q}(k-1)+k+1}$$

for some positive $\bar{C}$.

We next use Lemma 3 repeatedly and note that there are at least $\frac{1}{3}|S^n|$ valid ways to choose p. Moreover, given such a selection, there are at least $\binom{\lfloor \frac{n}{4} \rfloor}{k}$ ways to choose h and $\{c_i\}_{i=1}^{k-1}$. And given such a selection there are at least $\binom{\lfloor |S^n|/4 \rfloor}{q_0^h+(k-1)\bar{q}}$ valid selections of $(r_1, r_2, ...r_{q_0^h})$ and $\{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}$.

To complete the proof, we note that for different selections of

$$\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$$

the events $E^n\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ are disjoint. Thus, the expected number of students who are matched in some stable allocation, but are unmatched in another, which is greater than the expected number of events $E^n\left(p, r_1, r_2, ...r_{q_0^h}, h, c, c_1, c_2, ...c_{k-1}, \{\{s_{c_i}^j\}_{j=1}^{\bar{q}}\}_{i=1}^{k-1}\right)$ that are realized, is bounded below by $\hat{C}n$ for some positive $\hat{C}$.

□

Next, we formalize the claim from the body of the paper that Theorem 2 generalizes to a broad class of distributions over students' preferences.

Proposition 4. *Let $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$ be a sequence of uniform random markets, and let $\{\tilde{\Gamma}'^n\}_{n=1}^{\infty}$ be another sequence of random markets that differs only in the distribution of students' preferences.*[33] *Then, Theorem 2 holds for $\{\tilde{\Gamma}'^n\}_{n=1}^{\infty}$*

[33]Below, we present a definition of a random market that makes restrictions on the distribution of students' preferences. With a slight abuse of notation, we use the same term but do not make any restrictions.

so long as there exists some $\bar{C} > 0$ and $\alpha \in (0,1)$ such that for sufficiently large n, at least a fraction α of the events

$$E^n(p, r_1, r_2, ...r_{q^h_{\bar{t}_h}}, h, c)$$

and

$$E^n\left(p, r_1, r_2, ...r_{q^h_0}, h, c, c_1, c_2, ...c_{k-1}, \{\{s^j_{c_i}\}^{\bar{q}}_{j=1}\}^{k-1}_{i=1}\right)$$

with valid selections of colleges are at least $\bar{C}$ times as likely in Γ'^n as they are in Γ^n.

Proof. Follows immediately from our proof of Theorem 2. □

Allowing the preference-generating process from the body of the paper to draw uniformly from all permutations (not just acceptable ones) satisfies the conditions of the proposition, and so does having k, the number of colleges acceptable to each student, be student specific. The conditions also hold in sufficiently thick regular sequences of random markets (defined below), a broad family of sequences that generalizes the structure studied in Kojima and Pathak (2009) to the college admissions setting.

Let a *random market* be a tuple $\tilde{\Gamma} = \left\langle S, C, T, \{\succ_c, \gg_c, q^c_t\}^{c\in C}_{t\in T}, k, \kappa, \mathcal{D}\right\rangle$, where k is an integer greater than one, κ is a number in $(0,1)$, $\succ_c$ represents college c's strict preferences (which must be consistent with its master list, $\gg_c$, and list of quotas, $\{q^c_t\}_{t\in T}$), and $\mathcal{D} := \{p_c\}_{c\in C}$ is a distribution on C. A random market induces a college admissions market by drawing, for each student independently at random, preferences in the following way.[34]

- Step 0: Set $t = 1$, $A = \emptyset$, $B = \emptyset$, and let R be an empty ROL.
- Step $t \leq k|T|$: With probability κ proceed to step t.1. Otherwise, proceed to step t.2.
- Step t.1: If there are contracts with k different colleges on R, stop and set R as the student's preferences, where other contracts are not acceptable.[35] Otherwise draw a college, c, according to $\mathcal{D}$. If c is in $A \cup B$, repeat. Otherwise, append the ROL with the most generous contract with c, add c to B, and continue to step $t + 1$.

[34]Each randomization in the algorithm is independent.
[35]Alternatively, proceed to step t.2, and stop when R is longer than $k|T|$.

- Step $t.2$: If B is empty, continue to step $t.1$. Draw uniformly at random a college from B, $c \in B$. Append the ROL with the most generous contract with this college that does not appear on R. If the terms of this contract are the lowest in T, remove c from B and add it to A. Continue to step $t+1$.

A sequence of random markets, denoted by $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$, is *regular* if there exist integers $k, l, \bar{q},$ and λ, all greater than one, and $\kappa \in (0,1)$ such that:

1. $|C^n| = n$ for all n,
2. $k^n = k$, $\kappa^n = \kappa$, and $T^n = \{0, 1, ...l-1\}$ for all n,
3. $q_0^c \leq \bar{q}$ for all $c \in C^n$ and all n,
4. for all n, $c \in C^n$, and $s \in S^n$, $s \gg_c \emptyset$,
5. for all n and $c \in C^n$, there exist $t, t' \in T^n$ such that $q_t^c > q_{t'}^c > 0$, and
6. $\frac{1}{\lambda} n \leq |S^n| \leq \lambda n$, for all n.

A regular sequence of random markets, $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$, is *sufficiently thick* if there exist $\rho > 0$, $\omega \in (0,1)$, and an integer n' such that for all $n > n'$,

$$\frac{\max\limits_{c \in C^n} p_c^n}{\max\limits_{c \in C^n}{}_{\lceil \omega n \rceil} p_c^n} < \rho,$$

where $\max_i$ is the i-th highest element in a set. This condition means that the most popular college is at most ρ times as popular as the $\omega \times 100^{\text{th}} -$ tile college (i.e., the ratio of popularities does not grow without bound).

C Evidence from the IPMM

We complement the analysis from Section 4 by studying data from the Israeli Psychology Master's Match. We provide evidence in support of our assumptions on the demand structure, and corroborate the predictions of our model.

Background

In Israel, admissions to graduate programs in psychology are highly selective. Each year, about 1,400 students graduate from a bachelor's program, but this does not certify them to serve as therapists. In order to become a therapist, one needs to complete a clinical graduate degree and later an apprenticeship. But seats in clinical programs are scarce: only 300 students are accepted each year and about 300 students are accepted to other, non-clinical, programs.

While the number of applicants – approximately 1,000 a year – far exceeds the number of available seats, departments of psychology still compete for top talent. In an attempt to attract "star" applicants, several departments offer a limited number of prestigious scholarships to selected students. In 2014, when this market was centralized, it was critical to allow these departments to continue to pursue this recruiting strategy. Thus, the version of DA that is used in this market allows programs to offer contracts with multiple funding levels, and allows applicants to rank these alternatives separately.[36]

Data

We use the dataset prepared by Hassidim, Romm and Shorrer (2016). The data includes administrative match data, including ROLs and program reports from the 2014 and 2015 matches. Additionally, it includes the computer code for the market-clearing algorithm.

In 2014, ten programs in three departments offered admission under multiple funding levels. The number increased to 15 programs in four departments in 2015. Funding levels ranged from approximately \$2,000 a year to approximately \$20,000 a year. The number of available scholarships was 25 in 2014, and 36 in 2015.

Each year about 1,000 students participated in the match. The number of ROLs ranking some contract with a program offering multiple funding levels was 271 in 2014, and grew to 458 in 2015, as a result of the growth in the number of programs offering admission with multiple levels of funding (i.e., ignoring observations attributed to these programs only, the number remained almost constant).

[36]The mechanism used by the Israeli Psychology Master's Match is strategy-proof for students. For more details about the market and the mechanism, see Hassidim, Romm and Shorrer (2017).

Student rank-ordered lists

Participants ranked, on average, 4.32 contracts ($\sigma = 4.14$). About 37.2% of the participants ranked at least one of the programs that offered admission under multiple financial-aid levels. Of these, only 3.4% ranked only the funded contract in some program, but not the unfunded contract. Among the applicants who ranked both a funded and an unfunded contract, more than 90% ranked the funded contract first.[37] Among these applicants, the mean number of contracts ranked between a funded contract and the unfunded contract with the same program was 0.34. In 82.6% of the cases, the number was zero.

The core

While we have access to all match data, there are still data limitations that do not allow us to calculate the core fully. The main issue is that only the parts of the departments' preferences that were required to calculate the outcome of DA were elicited. And departments, who typically offer several study programs, often have complex preferences. An additional issue is that we are not aware of an efficient way to calculate the core.

Instead, we take an approach similar to the one we used in the Hungarian dataset for detecting applicants over whom the program may be able to apply market power (MP_c). Next, we declare them as ineligible for funding in that program (changing the department's preference report) and rerun the match. Finally, we check if the match is stable with respect to true preferences (i.e., if the applicant and the department are part of a blocking coalition of the resulting match). To do so, we require the assumption that programs do not care directly about the identities of the recipients of financial aid, but only about the quality of the incoming cohort. Based on our discussions with department chairs and recruiting committees during the design of the centralized clearinghouse, we are very comfortable with this assumption.

We find that, with a handful of exceptions, programs have market power over the recipients of financial aid. And while they can reallocate funding among admitted students, applying market power will not improve the quality of their incoming cohort. Our findings are reminiscent of Claim 1 in Example 1, which is not surprising in light of the fact that in 82.6% of the

[37] We believe that other applicants made a mistake. For details, see Hassidim, Romm and Shorrer (2016).

cases the funded and unfunded contracts in the same program were ranked consecutively.[38]

[38] In the 2016 match, two out of the four schools that had offered multiple levels of funding decided to offer identical terms to all students admitted to the same program.