

# The Large Core of College Admission Markets: Theory and Evidence\*

Péter Biró      Avinatan Hassidim      Assaf Romm      Ran I. Shorrer  
Sándor Sóvágó<sup>†</sup>

## Abstract

We study stable allocations in college admissions markets where students can attend the same college under different financial terms. The deferred acceptance algorithm identifies a stable allocation where funding is allocated based on merit. While merit-based stable allocations assign the same students to college, non-merit-based stable allocations may differ in the number of students assigned to college. In large markets, this possibility requires heterogeneity in applicants' sensitivity to financial terms. In Hungary, where such heterogeneity is present, a non-merit-based stable allocation would increase the number of assigned applicants by 1.9%, and affect 8.3% of the applicants relative to any merit-based stable allocation. These findings contrast sharply with findings from the matching (without contracts) literature.

---

\*This version: August, 2022. We thank Yannai Gonczarowski, Jacob Leshno, Bobby Pakzad-Hurson, Deborah Marciano, Thayer Morrill, Alex Rees-Jones, Al Roth, Ben Roth, Jan Christoph Schlegel, Alex Teytelboym and seminars and conferences participants for helpful discussions and suggestions. We are especially grateful to Scott Kominers, who discussed an earlier version of this paper in the 2018 ASSA meeting, for many useful suggestions. This research was supported by a grant from the United States – Israel Binational Science Foundation (BSF), Jerusalem, Israel. This research was supported by the Israel Science Foundation (grants no. 1241/12 and 1780/16). Assaf Romm is supported by a Falk Institute grant. The Hungarian Higher Education Application Database (FELVI) is owned by the Hungarian Education Bureau (Oktatási Hivatal). The data were processed by the Databank of KRTK (Centre for Economic and Regional Studies).

<sup>†</sup>Biró: Center for Economic and Regional Studies–Economics Institute, Corvinus University of Budapest (biro.peter@krtk.mta.hu); Hassidim: Bar Ilan University (avinatan@macs.biu.ac.il); Romm: Hebrew University of Jerusalem (assafr@gmail.com); Shorrer: Pennsylvania State University (rshorrer@gmail.com); Sóvágó: University of Groningen (s.sovago@rug.nl).

# 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 (Roth and Xing, 1994; Roth, 2002).<sup>1</sup> The matching market design literature finds that a designer who wishes to implement a stable allocation has limited scope for further design. First, the rural hospital theorem determines that the same positions are filled in all stable allocations (Roth, 1984a, 1986). Second, the set of stable allocations has the consensus property: all students prefer the outcome of the student-proposing deferred acceptance mechanism (henceforth SP-DA) to any other stable allocation (Gale and Shapley, 1962; Roth, 1984b). Third, empirical and theoretical studies suggest that all students, save for a handful, receive the same assignment in all stable allocations (e.g., Ashlagi et al., 2017; Azevedo and Leshno, 2016; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Roth and Peranson, 1999). This last finding implies that schools have limited incentive to collect information and to misreport their preferences (Demange et al., 1987).

The above-mentioned results apply to two-sided matching markets (men and women, students and schools, etc.) where agents' preferences are over potential partners from the other side. However, the environments studied and designed by economists are often more complex. For example, college applicants care not only about the study program they are assigned to, but also about the level of financial aid they receive. In this paper, we ask whether the set of stable allocations continues to be small in these more complex environments.

We study Hungarian college admissions, where colleges offer multiple levels of financial aid. In this environment, we show theoretically and empirically that it is no longer true that almost all students receive the same assignment in all stable allocations. Furthermore, different stable allocations differ in the set of positions filled and in the composition of assigned students, and there is no consensus among students on the most-preferred stable allocation.

To gain intuition, we provide a simple example. The example is a minimal instance of our model of Hungarian college admissions.<sup>2</sup> We accompany the verbal description with the notation of our model, which we introduce in Section 3.

**Example 1.** There are two students,  $S = \{r, p\}$ , and one college,  $C = \{c\}$ . The college has two seats, but only one of these seats is state-funded (formally,  $q_c^0 = 1, q_c^1 = 1$ ). The college finds the rich student,  $r$ , more attractive than the poor student,  $p$  (formally,  $r \gg_c p$ ). The college prefers to accept the most attractive students, and to fill its capacity. The rest of the college's preferences over acceptable allocations are fully described by

$$\{(r, c, 0), (p, c, 1)\} \succ_c \{(r, c, 1), (p, c, 0)\} \succ_c \{(r, c, 0)\} \succ_c \{(r, c, 1)\} \succ_c \{(p, c, 0)\} \succ_c \{(p, c, 1)\},$$

where 1 (0) indicates admission with (without) state funding.

<sup>1</sup>Stability is also useful for predicting behavior in decentralized matching markets (e.g., Banerjee et al., 2013).

<sup>2</sup>Our model captures the structure of preferences observed in other centralized college admissions markets, including Turkey, Australia, Israel, Ukraine, Russia, and the US.

The rich student,  $r$ , prefers to receive state funding, but she is willing to attend the college even if she does not get state funding. Formally,  $r$ 's preferences are  $(r, c, 1) \succ_r (r, c, 0) \succ_r \emptyset$ . By contrast, the poor student,  $p$ , is only interested in admission with state funding. Thus,  $p$ 's preferences are summarized by  $(p, c, 1) \succ_p \emptyset$ .

There are two stable allocations in this market: the result of SP-DA,  $\{(r, c, 1)\}$ , where funding is allocated based on merit; and  $\{(r, c, 0), (p, c, 1)\}$ , where more positions are filled and funding is not allocated based on merit. We note that the only allocation that both students weakly prefer to both of these stable allocations has both students receiving state funding, but this allocation is not acceptable to the college. Therefore, there is no consensus among students on a student-optimal stable allocation.

Can we expect stable allocations to be meaningfully different in realistically-sized markets? To answer this question, we develop a large market model of Hungarian college admissions markets and study the assumptions required from this model to guarantee that the set of stable allocations is large. Our main theoretical result (Theorem 3) is that the set of stable allocations is large when students are heterogeneous in their sensitivity to financial terms (like the rich student and the poor student in Example 1). By contrast, when students are not heterogeneous in their sensitivity to financial terms—that is, when all students find both contracts with the same college to be nearly perfect substitutes or when all students do not find them to be close substitutes—then the set of stable allocations is small. We also show that both variants of DA allocate funding based on merit (Theorem 1) and that the set of merit-based allocations is small.

The proof of Theorem 3 as well as our empirical analysis rely on the *preference flip algorithm*, a novel stable algorithm. To gain intuition on how our algorithm works, we revisit Example 1. If each type of seat was controlled separately, then competition between the seats would rule out the stable allocation in which both students are assigned. However, since one college controls both seats, it can make sure that they do not compete with each other. Given a stable allocation, colleges have *local market power* over students who are admitted with state funding, but who have no outside option (a contract at another college or being unassigned) that they prefer to the self-funded contract with the same college. An extreme case is when students rank the state-funded and self-funded contracts with the same college consecutively, like the rich student in our example. In this case, colleges can exercise market power by refusing to accept such students to state-funded seats, thus freeing up the state-funded seats which can then be used to recruit price-sensitive students like the poor student in our example.<sup>3</sup> The preference flip algorithm can be interpreted as allowing colleges to exercise their local market power.

We use the preference flip algorithm for exploring the size of the set of stable allocations. However, implementing it will require more than merely changing the matching algorithm. Specifically, the preference flip algorithm requires colleges to offer financial aid selectively based on students' sensitivity to financial terms, giving students an incentive to misrepresent their sensitivity to financial terms. Still, colleges can use hard information to infer students' sensitivity to financial terms. For example, in the American context, such information is available to colleges through the Free Application for Federal Student Aid (FAFSA)

<sup>3</sup>This argument is not precise, since by recruiting new students the college may change the outside options available to its other students. Our analysis takes this challenge into account.

form. In addition, many American colleges hire companies to personalize their financial aid packages using large proprietary datasets (Lieber, 2021; Caniglia and Porterfield, forthcoming).

We next assess the size of the set of stable allocations empirically using administrative data from the Hungarian college admissions system. We document substantial heterogeneity in applicants' sensitivity to financial terms. In line with our large market analysis, both variants of DA result in essentially identical allocations, while the outcome of the preference flip algorithm differs for more than 9,000 applicants (approximately 8 percent of the applicants), with slightly more winners than losers. More importantly, the preference flip algorithm increases the number of applicants assigned to college by 1.9 percent (approximately 1,500) with approximately 2,100 unassigned applicants gaining admission and approximately 600 losing their place.<sup>4</sup> Since programs in the capital—which are generally prestigious and highly demanded—typically fill all their positions both under DA and under our alternative, the gains in enrollment mostly accrue to colleges in the periphery.

We analyze the characteristics of those who benefit and lose from the preference flip algorithm. Applicants who prefer the outcome of the non-merit-based preference flip algorithm to the outcome of DA come from lower socioeconomic background relative to applicants who prefer the outcome of DA. The preference flip algorithm also increases geographic mobility by increasing the number of applicants assigned to a college outside their county of residence.

Our data does not include hard information that would allow inferring applicants' sensitivity to financial terms. For this reason, our main analysis uses an algorithm which only applies local market power over students who are extremely insensitive to financial aid (as revealed by their preference reports). Had hard information been collected, one can argue that these students could plausibly be identified. In Appendix D, we expand the analysis in two directions. First, we assess the consequences of being able to detect only a subset of these students. Second, we analyze algorithms that apply market power more broadly. We find that the effectiveness of our approach increases with the quality of information on students' sensitivity to funding.

Our findings speak to the broader research question of how to design college admissions processes in the presence of constraints on additional resources such as financial aid, dormitories, etc. Our analysis does not consider many policy levers that are available to policymakers. For example, we fix the number of scholarships the government allocates to each program, we keep scholarships indivisible, and we require that admission and financial terms be determined simultaneously based on preference reports and that the outcome is stable. Even under these restrictions, we find a substantial scope for market design to affect the outcome.

The remainder of the paper is organized as follows. Following a short review of related literature, Section 2 describes college admissions in Hungary. Section 3 introduces a formal

---

<sup>4</sup>Having found that the set of stable allocations could be large and in the absence of the consensus property, a natural objective is to identify stable allocations that are optimal according to some criterion. However, we show that achieving this objective may not be feasible with current computing technology. Specifically, Theorem 2 shows that finding the largest stable allocation is NP-hard. This means that finding an efficient (polynomial-time) algorithm for solving this problem will establish that  $P = NP$  (which is widely believed to be false).

model of Hungarian college admissions. Section 4 explores the set of stable allocations. Section 5 includes the large market analysis. Section 6 reviews the empirical findings. Section 7 concludes. Proofs are relegated to the appendix throughout.

## 1.1 Related Literature

Empirical and theoretical studies show that the set of stable allocations is typically small (Ashlagi et al., 2017; Azevedo and Leshno, 2016; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Roth and Peranson, 1999; Storms, 2013). Prior to these studies, a large portion of the literature on the theory of two-sided matching markets was motivated by the potential multiplicity of stable allocations. Examples include studies of the structure of the set of stable allocations (Knuth, 1976), of fair stable allocations (Klaus and Klijn, 2006; Schwarz and Yenmez, 2011), and of incentives (Ehlers and Massó, 2007; Roth, 1982; Sönmez, 1999). The rural hospital theorem (Roth, 1984a, 1986) refuted suggestions that changing the way the National Residency Match Program treats medical graduates and hospitals may change the number of doctors assigned to rural hospitals.

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). Demange et al. (1987) show that schools' incentives to manipulate SP-DA are intimately related to the multiplicity of stable allocations.<sup>5</sup> Several studies show that in large markets it is safe for schools to report their true preferences to SP-DA (Ashlagi et al., 2017; Azevedo and Budish, 2018; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009; Lee, 2017).

Complex two-sided matching markets are studied in the matching-with-contracts literature (Hatfield and Milgrom, 2005). Much of this literature focuses on identifying conditions under which SP-DA remains stable and strategy-proof for students. Examples include Fleiner (2003), Hatfield and Kojima (2010), Hatfield and Milgrom (2005), Hatfield et al. (2020), Kelso and Crawford (1982), and Roth (1984b).<sup>6</sup> Most closely related is Hatfield and Kominers (2015).

Hassidim et al. (2017a) use the results of Hatfield and Kominers (2015) to show that SP-DA is stable and strategy-proof in more general college admissions environments. They show that different stable allocations may have different cardinality. The proofs in our paper rely on a different construction, which allows us to derive further results. We contribute to Hatfield and Kominers (2015) and Hatfield et al. (2020) in several ways. First, we identify a real-life market design application where colleges' choice functions satisfy the hidden substitutes condition but fall outside the (unilateral) substitutes domain.<sup>7</sup> Second, we provide

---

<sup>5</sup>Numerous studies have analyzed the optimal behavior of schools when the SP-DA mechanism is in place (e.g., Coles and Shorrer, 2014; Ehlers, 2004; Konishi and Ünver, 2006; Roth and Rothblum, 1999; Sönmez, 1997).

<sup>6</sup>Applications related to college admissions include Abizada (2016), Afacan (2020), Aygün and Bó (2016), Aygün and Turhan (2017), Nei and Pakzad-Hurson (2016), Pakzad-Hurson (2014), Westkamp (2013), and Yenmez (2018).

<sup>7</sup>To see that college preferences are not (unilaterally) substitutable (Hatfield and Kojima, 2010), note that in Example 1, college  $c$  rejects  $(p, c, 1)$  from the menu  $\{(p, c, 1), (r, c, 1)\}$  but does not reject this contract from the larger menu  $\{(p, c, 1), (r, c, 1), (r, c, 0)\}$ . Colleges' preferences also do not meet any other condition that

a large market analysis. Third, we provide a complete characterization of the set of stable allocations in the domain of Hungarian college admissions markets.<sup>8</sup> Fourth, we show that different stable allocations may have different cardinality in this more restrictive domain, and that stable allocations in this domain may result in fewer assigned students relative to SP-DA. Fifth, we consider the practical aspect of computation.

Large market models have been studied extensively in the market design literature. Examples include [Ashlagi and Nikzad \(2020\)](#), [Ashlagi et al. \(2014, 2019, 2017\)](#), [Azevedo and Leshno \(2016\)](#), [Immorlica et al. \(2020\)](#), [Kojima and Pathak \(2009\)](#), [Kojima et al. \(2013\)](#), [Lee \(2017\)](#), and [Leshno and Lo \(2020\)](#). Most closely related are [Immorlica and Mahdian \(2005, 2015\)](#) and [Rheingans-Yoo \(2020\)](#) who show that even in the absence of contracts, certain regularities in preferences can lead to a large set of stable allocations. These papers study settings where the rural hospital theorem and the consensus property apply. By contrast, we study markets with contracts and show that different stable allocations may differ in the number of students assigned to college and where the consensus property does not hold.

Our paper is related to studies of reserve design ([Aygün and Turhan, 2017](#); [Dur et al., 2018](#); [Aygün and Turhan, 2020](#); [Pathak et al., 2020](#)). In the context of school choice, a key observation in this literature is that students may be indifferent between different seats in the same school, but some seats are reserved for certain groups of students. This implies, using our terminology, that schools have local market power over students from these groups. The reserve-design literature focuses on the effect of different ways the mechanism can break students' preference ties in order 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.

Our results shed light on the policy debate on market power in higher education (see, e.g., [Hoxby, 2000](#)). This literature gained traction after the U.S. Department of Justice brought 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 MIT does 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 findings provide support to MIT's arguments. We show that even in the absence of a motive to increase profit, colleges 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. On average, needy students gain, in part at the expense of more wealthy students. Furthermore, information about students' sensitivity to financial aid, is necessary to facilitate such behavior.

---

guarantees the existence of a student-optimal stable allocation ([Hassidim et al., 2019](#)).

<sup>8</sup>The substitutable completions approach of [Hatfield and Kominers](#) guarantees existence, but it is not guaranteed to identify all stable allocations. Indeed, one can verify that, in [Example 1](#), for any substitutable completion of the college's choice function,  $Y^{SP-DA}$  is the unique allocation that is stable with respect to the completion.



## 2 Background: College Admissions in Hungary

Each year about 80,000 applicants are assigned to undergraduate college programs in Hungary. Applicants are assigned to *programs*, i.e., a specific major in a specific college.<sup>9</sup> The admissions process has been centralized since 1985. The centralized clearinghouse has been using variants of DA to assign applicants to colleges since 1997.

Historically, higher education in Hungary was free. However, in recent years, the government has capped the number of state-funded (free) seats, and programs are allowed to offer admission to self-funded (tuition-paying) seats as well.<sup>10</sup> To implement this change, the centralized clearinghouse started requesting applicants to submit a rank-order list (ROL) of *alternatives*, specifying that the student attends a program under particular financial terms (e.g., economics in the University of Debrecen with state funding). This flexibility was crucial in order to allow applicants to express their preferences. For example, some applicants may not be willing or able to afford paying tuition, while others may have a strong preference for certain programs, and be less price-sensitive.

Applicants may rank as many alternatives as they wish. Submitting an ROL with three programs (corresponding to up to six alternatives) only requires paying the fixed application fee of \$50 (9,000HUF). Applicants are required to pay a registration fee of \$11 (2,000HUF) for each additional program in their ROL. Disadvantaged applicants are exempt from these fees.<sup>11</sup>

College programs report to the mechanism their capacity (i.e., the maximal number of applicants they can accept). Admissions priorities in each program are based on a weighted average of several variables (mainly academic performance in the 11th and 12th grades and matriculation exam scores, but also credits for disadvantaged and disabled applicants, for applicants who demonstrate fluency in another language, and for a small number of gifted applicants). Different study programs may use different weighting schemes (e.g., a computer science program may assign a greater weight to physics grades relative to a psychology program).

In 2007, the year we focus on, to accommodate the multiple modes of financing, the mechanism required that the programs also report the number of state-funded seats they offered. This took place prior to the beginning of the admissions process.<sup>12</sup> The mechanism then created two “auxiliary programs” with capacities corresponding to the number of state-funded and self-funded seats in the program, and endowed both with the priorities of the program.<sup>13</sup> Finally, student-receiving DA (SR-DA)—i.e., auxiliary-program-proposing

---

<sup>9</sup>We focus on admissions to undergraduate programs, which include three types of programs: bachelor’s degree programs (with a typical study duration of 3 years), combined bachelor–master’s programs (with a typical study duration of 5 years), and tertiary vocational education (with a typical study duration of 2 years).

<sup>10</sup>Citizens of the European Economic Area who have not yet graduated from higher education are eligible for state funding.

<sup>11</sup>To be eligible for disadvantaged status, an applicant must have a per-capita household income that is lower than 130 percent of the minimum pension (i.e., lower than approximately \$1,500 a year).

<sup>12</sup>In 2008, Hungarian college admissions adopted a more complex variant of DA (Biró et al., 2010; Shorrer and Sóvágó, 2018). Although our analysis continues to hold under the new mechanism, we focus on 2007 to highlight that our results do not rely on the new, unique features.

<sup>13</sup>This description applies to full-time programs. There was an additional constraint on the total number of

DA—was used to determine the final assignment.

Table 1 presents information on the alternatives that colleges offered in 2007. The applicants' choice set consisted of 3,740 alternatives, corresponding to 2,289 study programs. The number of study programs that offered both state-funded and self-funded alternatives was 1,451. Some study programs were available with state-funding exclusively (172), and other programs—mostly part-time programs—were available with self-funding exclusively (666). Panel B shows that 93,999 of the 111,685 available seats (84.1 percent) were available in study programs that offered both state-funded and self-funded alternatives. The overwhelming majority of the seats that were offered with self-funding exclusively (about 91 percent) were in the form of part-time education.

Table 1: Applicants' choice set

	Total (1)	Full-time (2)	Part-time (3)
<i>A. Alternatives</i>			
Number of programs	2,289	1,357	932
– state- and self-funded	1,451	1,133	318
– state-funded exclusively	172	158	14
– self-funded exclusively	666	66	600
Number of alternatives	3,740	2,490	1,250
<i>B. Capacities</i>			
State-funded and self-funded	93,999	75,197	18,802
– state-funded	47,809	43,084	4,725
– self-funded	46,190	32,113	14,077
State-funded exclusively	917	806	111
Self-funded exclusively	16,769	691	16,078
Total	111,685	76,694	34,991

*Notes:* The table presents summary statistics of the applicants' choice set. Panel A presents the number of alternatives by the financial terms. Panel B displays the number of seats by the financial terms.

**Timeline.** The application process proceeds as follows. First, the centralized clearinghouse publishes a booklet that includes the rules of the college admissions process together with the list of the alternatives (program–financial terms pairs), and the number of seats available in each alternative in November. Applicants submit their ROLs in mid-February. In May and June, applicants take their matriculation exam. Finally, in mid-July, the clearinghouse notifies applicants about their placement and publishes the priority-score cutoffs for each alternative, i.e., the minimum priority score that was needed to gain admission.

state-funded seats in each field of study, which was only binding for part-time programs. Additionally, part-time programs in computer science and engineering use slightly different priorities for state-funded and self-funded seats (they assign a lower weight to mathematics scores of applicants for self-funded seats relative to state-funded seats). Our empirical analysis takes this into account (see Appendix E).



**Financing Colleges.** The government funds colleges on a per-student basis, irrespective of the financial terms. Additionally, colleges collect tuition from self-funded students,<sup>14</sup> and receive additional compensation for state-funded students from the government. Thus, holding the number of state-funded (self-funded) students constant, colleges' revenue increases with each additional admitted self-funded (state-funded) student.

### 3 A Model of Hungarian College Admissions

We use the many-to-one matching-with-contracts model of [Hatfield and Milgrom \(2005\)](#) to describe *Hungarian college admissions markets*. There is a finite set of *colleges*,  $C$ , a finite set of *students*,  $S$ , and a set of *financial terms*,  $T = \{0, 1\}$ . A *contract* is a tuple  $(s, c, t) \in S \times C \times T$  that specifies a student, a college, and financial terms (with  $t = 1$  representing state funding, and  $t = 0$  representing self funding).

An *allocation* is a subset  $Y \subseteq S \times C \times T$ . An allocation  $Y$  is *feasible* if no student is included in more than one contract. Formally, for each student  $s$ ,  $|Y \cap (\{s\} \times C \times T)| \leq 1$ . Given an allocation,  $Y$ , we let  $Y_S$  denote the set of students involved in some contract in  $Y$ . Formally,  $Y_S := \{s \in S \mid Y \cap (\{s\} \times C \times T) \neq \emptyset\}$ . Similarly,  $Y_C := \{c \in C \mid Y \cap (S \times \{c\} \times T) \neq \emptyset\}$ .

**Students' preferences** Each student,  $s$ , has strict preferences over contracts in  $\{s\} \times C \times T$  (not all of them are necessarily available) and an outside option which we denote by  $\emptyset$ . We denote student  $s$ 's preference relation by  $\succ_s$ . Students' preferences, therefore, induce weak preferences over all feasible allocations, where students only consider their own assignment.

**Colleges' preferences** Each college,  $c$ , has strict preferences,  $\succ_c$ , over all feasible allocations in  $S \times \{c\} \times T$ .<sup>15</sup> These preferences induce weak preferences over all feasible allocations, where  $c$  strictly prefers  $Y$  to  $Y'$  if and only if  $Y \cap (S \times \{c\} \times T) \succ_c Y' \cap (S \times \{c\} \times T)$ . The preferences of each college,  $c$ , satisfy the following conditions.

First,  $c$  is associated with two numbers,  $q_c^1$  and  $q_c^0$ , representing a constraint on the number of students that can be accepted under each of the financial terms. The college  $c$  prefers the empty allocation to all allocations that *violate  $c$ 's quotas*, that is, that assign to  $c$  more than  $q_c^t$  students under the financial terms  $t$ .

Second,  $c$  has a complete order over  $S \cup \{\emptyset\}$ , denoted by  $\gg_c$ , representing a ranking over students. Given an allocation  $Y \subseteq S \times \{c\} \times T$  (an assignment of students to the college under some financial terms), if  $q_c^t$  is not binding then the college prefers to accept an additional student  $s$  under the financial terms  $t$  if and only if  $s \gg_c \emptyset$ . Formally, if  $s \notin Y_S$  and  $|Y \cap (S \times \{c\} \times \{t\})| < q_c^t$ , then  $Y \cup \{(s, c, t)\} \succ_c Y$  iff  $s \gg_c \emptyset$ . Additionally,  $c$  prefers to replace student  $s$  who receives the financial terms  $t$  with another student  $s'$  who is not assigned to  $c$  (under the same financial terms) if and only if  $s' \gg_c s$ . Formally, for all  $Y \subseteq S \times \{c\} \times T$ , if  $(s, c, t) \in Y$  and  $s' \notin Y_S$ , then  $(Y \cup \{(s', c, t)\}) \setminus \{(s, c, t)\} \succ_c Y$  iff  $s' \gg_c s$ .

Third, so long as quotas are not violated, the composition of the incoming cohort is lexicographically more important to the college relative to the way funding is allocated. Formally, let  $Y, Y' \subseteq S \times \{c\} \times T$  be two feasible allocations that do not violate  $c$ 's quotas. Then

<sup>14</sup>Tuition for a 3-year bachelor's degree program ranges from \$3,280 to \$6,560 (600,000–1,200,000HUF).

<sup>15</sup>None of our results relies on the assumption that colleges' preferences are strict.

if  $Y_S = Y'_S$  (i.e.,  $Y$  and  $Y'$  differ only in the identity of the recipients of state funding in  $c$ ) and  $Y \succ_c Y'' \succ_c Y'$  for some  $Y'' \subseteq S \times \{c\} \times T$ , it follows that  $Y''_S = Y_S$ .

**Discussion of theoretical assumptions** We refrain from assuming that students always prefer to receive state funding for several reasons. First, this generality allows our model to capture environments where the ranking of alternatives is more ambiguous, such as admission with and without a dormitory assignment. Second, in practice, some students are not eligible to receive financial aid and a small fraction of applicants report unusual preferences (Shorrer and Sóvágó, 2018).

According to our assumptions, subject to quotas, colleges' preferences depend more on the size and quality of the incoming cohort than they do on the distribution of funding. This assumption is not unusual in the literature (e.g., Kim, 2010; Heo, 2017), and is consistent with preferences of colleges in other markets (e.g., Hassidim et al., 2017b). Our assumptions rule out redistributive motives. As highlighted by Caniglia and Porterfield (forthcoming), the presence of such motives will strengthen our arguments.

Our model also abstracts from upper quotas on state-funded seats in part-time programs in the same field (Biró et al., 2010). We note that all of our theoretical results generalize to this more complex environment. The upper quotas create links between part-time programs, making their choice functions more complex. The added complexity creates further opportunities to apply local market power (e.g., transfer students between part-time programs). We choose not to take advantage of this feature, which is specific to the Hungarian market, in our empirical analysis.

The special case of our model where  $\min\{q_c^0, q_c^1\} = 0$  for all  $c$  corresponds to *two-sided many-to-one matching (without contracts) markets with responsive preferences* (Roth, 1985). Given such a market, we often refer to college  $c$ 's only positive quota as  $c$ 's *quota* and to allocations as *matchings*.

**Individual rationality** The preferences of each college  $c$  induce a choice function  $\text{Ch}_c : 2^{S \times C \times T} \rightarrow 2^{S \times \{c\} \times T}$ , that identifies the feasible subset of  $Y \cap (S \times \{c\} \times T)$  most preferred by  $c$ . Similarly, for students,  $\text{Ch}_s : 2^{S \times C \times T} \rightarrow 2^{\{s\} \times C \times T}$  chooses the most preferred acceptable contract involving  $s$  (if one exists). An allocation  $Y$  is *individually rational* if all agents choose all the contracts in which they are involved, that is  $\bigcup_{c \in C} \text{Ch}_c(Y) = \bigcup_{s \in S} \text{Ch}_s(Y) = Y$ . Individually rational allocations cannot violate quotas since colleges prefer the empty set to allocations that violate their quotas. Furthermore, they are feasible since students never choose more than one contract ( $|\text{Ch}_s(\cdot)| \leq 1$  for all  $s$ ).

**Stability** An allocation  $Y$  is *blocked (through  $Z$ )* if there exists a college,  $c$ , and a non-empty set  $Z \subseteq (S \times \{c\} \times T) \setminus Y$  such that  $Z \subseteq \text{Ch}_c(Y \cup Z)$  and  $Z \subseteq \bigcup_{s \in S} \text{Ch}_s(Y \cup Z)$ . An allocation is *stable* if it is individually rational and not blocked. In words, an allocation is stable if no coalition of agents can achieve a weak improvement on its own.

The following proposition provides a complete characterization of stable allocations.

**Proposition 1.** *An allocation  $Y$  is stable if and only if all of the following hold:*

1.  $Y$  is individually rational.

2.  $Y$  is not blocked through a singleton  $\{(s, c, t)\}$ .
3.  $Y$  is not blocked through a change in the financial terms of one student, and a contract with a new student using the freed-up space. Formally, for all  $(s, c, 1 - t) \notin Y$  and  $(s', c, t) \in Y$ ,  $Y$  is not blocked through  $\{(s', c, 1 - t), (s, c, t)\}$ .
4.  $Y$  is not blocked through a change in financial terms that keeps students in the same college (but under different financial terms). Formally,  $Y$  is not blocked through  $Z$  such that  $(s, c, t) \in Z$  implies  $(s', c, 1 - t) \in Y$ .

We note that Conditions 3 and 4 of Proposition 1 can only be violated in the presence of multiple contractual terms. Therefore, in two-sided many-to-one matching (without-contracts, i.e., when  $|T| = 1$ ) markets with responsive preferences the proposition reduces to the statement that a matching is stable if and only if it is individually rational and not blocked by a student–college pair.

**Certainly stable allocations** Allocation  $Y$  is *certainly stable* if Conditions 1–3 of Proposition 1 hold, as well as the following condition:

- 4'. For each  $(s, c, t) \in Y$ , if  $(s, c, 1 - t) \succ_s (s, c, t)$  then  $|Y \cap (S \times \{c\} \times \{1 - t\})| = q_c^{1-t}$  and  $(s', c, 1 - t) \succ_{s'} (s', c, t)$  for all  $(s', c, 1 - t) \in Y$ .

When Conditions 1-3 of Proposition 1 hold, Condition 4' implies Condition 4 of Proposition 1. Therefore, a certainly stable allocation is also stable.<sup>16</sup> The notion of certain stability proves useful in our empirical analysis because it only relies on quotas and rankings (i.e.,  $q_c^0$ ,  $q_c^1$ , and  $\succ_c$ ) but not on the full description of  $\succ_c$ , and because there are fewer conditions to verify (each condition refers to no more than three agents).

**Comparing stable allocations** When comparing the stable allocation  $Y'$  to the stable allocation  $Y$ , the set of *winners*,  $W(Y', Y)$ , consists of all students that prefer their assignment under  $Y'$ . The set of winners comprises of the following subsets:

1. The set of *newly assigned* students,  $N(Y', Y)$ , consists of students who are assigned under  $Y'$  but not under  $Y$ .
2. The set of *funding winners*,  $FW(Y', Y)$ , consists of students who are assigned to the same college in both allocations, but prefer their financial terms under  $Y'$ . If all students prefer to be funded, students in  $FW(Y', Y)$  receive state-funding under  $Y'$ , but not under  $Y$ .
3. The set of *upwards transfers*  $UT(Y', Y) := W(Y', Y) \setminus (N(Y', Y) \cup FW(Y', Y))$  consists of winners who get assigned to a new, preferred college.

We define the set of *losers*, and its subsets, *newly unassigned*, *funding losers*, and *downward transfers*, symmetrically.

<sup>16</sup>The converse is not true because students may want to swap their financial terms in the same college in a stable allocation, as long as the college prefers that they do not.

**Merit-based allocations** We say that an allocation  $Y$  *allocates funding based on merit* if  $s$  weakly prefers  $(s, c, t)$  to  $(s, c, t')$  whenever  $(s, c, t)$  and  $(s', c, t')$  belong to  $Y$  and  $s \gg_c s'$ . When  $Y$  allocates funding based on merit, within each college,  $c$ , higher-ranked students according to  $\gg_c$  do not envy the financial terms of lower-ranked students who are also assigned to  $c$ . Specifically, if all students prefer to be funded, the recipients of funding in each college are all ranked higher than all self-funded students. A stable allocation that allocates funding based on merit is a *merit-based stable allocation*.<sup>17</sup>

**Related matching markets** The *related (two-sided) matching market (without contracts)* is an auxiliary market where each program-funding-level pair plays the role of an individual college. This mathematical object often simplifies our exposition and analysis. Formally, given a Hungarian college admissions market,  $\langle S, C, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C}, \{\succ_s\}_{s \in S} \rangle$ , the *related matching market* comprises the set of *auxiliary colleges*  $\hat{C} \equiv C \times T$  with responsive preferences, where  $(c, t) \in \hat{C}$  has quota  $q_c^t$  and uses the ranking  $\gg_c$ , and the set of students  $S$  with preferences  $\{\hat{\succ}_s\}_{s \in S}$ , such that for every  $s \in S$ ,  $c \in C$ , and  $t \in T$ ,  $(c, t) \hat{\succ}_s (c', t')$  if and only if  $(s, c, t) \succ_s (s, c', t')$ . A feasible allocation  $Y$  and a matching (in the related matching market) are *corresponding* if  $s$  matched with  $(c, t)$  if and only if  $(s, c, t) \in Y$ .

**Deferred acceptance** Student-proposing deferred acceptance takes a profile of preferences as input and outputs an allocation. The allocation is the result of the following process. In each round, each student proposes the most preferred acceptable contract from which she has not yet been rejected, if such a contract exists. Each college then rejects all but the most preferred subset of contracts from these proposals. The algorithm terminates when no proposal is rejected, and the output is the set of contracts proposed in the last round. Given a Hungarian college admissions market, we denote the outcome of SP-DA by  $Y^{\text{SP-DA}}$ .

Student-receiving deferred acceptance outputs the allocation that corresponds to auxiliary-college-proposing deferred acceptance (formally, the allocation to which this matching corresponds). Of note, in each round of the algorithm, a college may make multiple offers to the same student (one by each auxiliary college). Given a Hungarian college admissions market, we denote the outcome of SR-DA by  $Y^{\text{SR-DA}}$ .

In two-sided many-to-one matching (without contracts) markets with responsive preferences, SR-DA is known as college-proposing DA. We intentionally avoid this label to highlight the fact that colleges are not necessarily proposing the most preferred subset of contracts that were not rejected. Since colleges' preferences are not substitutable (Hatfield and Milgrom, 2005), such a process would sometimes “renege” on proposals that were not rejected due to the rejection of complementary contracts.<sup>18</sup>

<sup>17</sup>Abdulkadiroğlu and Sönmez (2003) establish a connection between stability and justified-envy-freeness in what became the standard school choice framework. Romm et al. (2020) study the relation between stability and the elimination of justified envy in a general matching with contracts markets. In the context of multi-dimensional constraints, Delacrétaz et al. (2016) note that a weak form of no justified envy is independent of stability.

<sup>18</sup>To see that a “college proposing” process would sometimes “renege” on proposals, consider a modification to Example 1 where neither student finds the unfunded seat acceptable. In the first round, the poor student is offered a funded seat and the rich student is offered an unfunded seat, which she declines. As a result, in later rounds, the college would like to renege on the offer made to the poor student, and instead offer the

**Proposition 2.**  $Y^{SP-DA}$  corresponds to the student-optimal stable matching in the related market.

**Corollary 1** (Strategic properties of DA). In Hungarian college admissions markets:

1. SP-DA is strategy-proof for students.
2. A student can manipulate SR-DA if and only if she strictly prefers  $Y^{SP-DA}$  to  $Y^{SR-DA}$ .

Corollary 1 provides a bound on the number of applicants with incentives to misrepresent their preferences to SR-DA.

## 4 The Set of Stable Allocations

We begin in Section 4.1 by studying the set of merit-based stable allocations. We show that this set is nonempty, and that it includes the allocations that result from both SP-DA and SR-DA. We also show that each college recruits the same number of students across all merit-based stable allocations. In Section 4.2, we note that different stable allocations differ in the number of assigned students and that there does not necessarily exist a student-optimal stable allocation. We introduce the preference flip algorithm, a (non-merit-based) algorithm designed to increase the number of admitted students compared to SR-DA and prove that it is stable. We then demonstrate that achieving certain policy goals may not be feasible with current computing technology. Specifically, Theorem 2 states that finding a largest stable allocation is NP-hard.<sup>19</sup>

### 4.1 Merit-Based Stable Allocations

We begin by establishing a connection between merit-based stable allocations and stable allocations in the related matching market.

**Theorem 1.** *Given a Hungarian college admissions market:*

1.  $Y$  is a merit-based certainly stable allocation if and only if the corresponding matching is stable in the related matching market.
2. If  $Y$  is a merit-based stable allocation, then there exists a merit-based certainly stable allocation  $Y'$  that all students weakly prefer which assigns all students to the same colleges (potentially under different contractual terms).

**Corollary 2** (Existence). The set of merit-based stable allocations is nonempty in Hungarian college admissions markets. In particular, it includes  $Y^{SP-DA}$  and  $Y^{SR-DA}$ .

**Corollary 3** (Weak rural hospital theorem). In Hungarian college admissions markets:

1. The set of students assigned to some college is identical under all merit-based stable allocations. Formally,  $[Y^{SP-DA}]_S = [Y]_S$  for any merit-based stable allocation  $Y$ .

---

funded seat to the rich student.

<sup>19</sup>Recent advances in the market design literature show that solving some NP-hard problems is feasible in practice (see [Ágoston et al., 2016](#); [Ágoston et al., 2021](#); [Leyton-Brown et al., 2017](#)).



Moreover, if  $Y$  is a certainly stable merit-based allocation, then for each  $c \in C$  and  $t \in \{0, 1\}$ :

2. The number of state-funded (self-funded) students assigned to  $c$  is equal under  $Y^{\text{SP-DA}}$  and  $Y$ . Formally,  $|Y^{\text{SP-DA}} \cap (S \times \{c\} \times \{t\})| = |Y \cap (S \times \{c\} \times \{t\})|$ .
3. If college  $c$  does not fill one of its quotas under SP-DA, the same students are assigned to  $c$  with these financial terms under  $Y$ . Formally, if  $|Y^{\text{SP-DA}} \cap (S \times \{c\} \times \{t\})| < q_c^t$  then  $Y^{\text{SP-DA}} \cap (S \times \{c\} \times \{t\}) = Y \cap (S \times \{c\} \times \{t\})$ .

Taken together, Corollaries 2 and 3 show that changing the orientation of DA cannot lead to a change in the number or even the composition of the assigned students.

**Corollary 4** (Weak consensus property). In Hungarian college admissions markets, all students weakly prefer  $Y^{\text{SP-DA}}$  to any other merit-based stable allocation.

Taken together, Corollaries 2 and 4 imply that it is possible to identify the student-optimal merit-based stable allocation time-efficiently (using the time-efficient SP-DA).

## 4.2 Beyond Merit-Based Stable Allocations

We begin by showing that Hungarian college admissions markets may admit stable allocations where funding is not allocated based on merit, and that the rural hospital theorem and the consensus property do not extend to this domain.

**Proposition 3.** *In Hungarian college admissions markets:*

1. *There may exist non-merit-based stable allocations.*
2. *Students may disagree on the most preferred stable allocation.*
3. *Different stable allocations may have more or fewer assigned students than  $Y^{\text{SP-DA}}$ .*

Example 1 has already shown a non-merit-based stable allocation that assigns more students to college than the unique merit-based stable allocation. In the merit-based stable allocation, the higher-ranked (rich) student receives funding even though her preferences are not sensitive to financial terms while the lower-ranked (poor) student, who is sensitive to financial terms, remains unassigned. In the non-merit-based stable allocation, the college applies its local market power over the rich student; i.e., the college refuses to allocate state funding to the rich student who is not sensitive to financial terms.

Building on this intuition, we next present a stable algorithm designed to increase the number of admitted students compared to SR-DA. The algorithm relies on the application of local market power over students who are not sensitive to financial terms. We say that a student,  $s$ , is *not sensitive to financial terms in college  $c$*  if she considers contracts with  $c$  to be nearly perfect substitutes. Formally, for any  $c' \in C$  and any  $t \in T$ ,  $(s, c, t) \succ_s (s, c', t) \iff (s, c, 1 - t) \succ_s (s, c', t)$  and  $(s, c, 1 - t) \succ_s \emptyset \iff (s, c, t) \succ_s \emptyset$ . A student is *not sensitive to financial terms* if she is not sensitive to financial terms in any college.



**Preference Flip Algorithm** Initialize a set  $A' \subseteq A$ , where  $A$  is the set of student–college pairs such that the student is not sensitive to financial terms in the college. For each pair in  $A'$ , flip the order of the contracts with the college in the student’s original ROL, so that the state-funded contract appears immediately *after* the self-funded contract, and run SP-DA on the resulting problem. If the resulting allocation is certainly stable (with respect to the original preferences), stop and output this allocation. Otherwise, remove some pairs from  $A'$  and repeat the process.

The description of the preference flip algorithm intentionally leaves two degrees of freedom: the choice of  $A'$  and the choice of elements to remove from this set. In the empirical application, we implement the preference flip algorithm by initializing  $A' = A$  and by removing pairs corresponding to the highest-ranked student in some college such that the pair is in  $A'$  and the state-funded contract between them is part of a potentially blocking allocation as described in the definition of certainly stable allocations (or a random pair, if no such pair exists). We note that with this choice the preference flip algorithm is computationally efficient.

**Proposition 4.** *For any initial set  $A' \subseteq A$  and any rule regarding the choice of elements to remove from  $A'$ , the preference flip algorithm results in a stable allocation. Furthermore, the resulting allocation may not allocate funding based on merit.*

In the main analysis we use the preference flip algorithm, which only applies local market power over students who are not sensitive to financial terms. These students could plausibly be identified using hard information. Supplemental Appendix D explores algorithms that apply local market power more broadly. These algorithms use the observation that a college has market power over students whose best feasible alternative to the funded contract with the college is attending the college without state funding, even if they do not rank both contracts with the college consecutively. Identifying these students in practice would pose a bigger challenge.

In exploring the set of stable allocations, a natural objective is to identify a stable allocation that assigns the largest number of students. We conclude this section by showing that this objective may not be feasible with current computation technology when the number of students and colleges is large.

**Theorem 2.** *Finding a maximum-size stable (or certainly stable) allocation in Hungarian college admissions markets is NP-hard.*

Theorem 2 determines that finding a maximum-size stable allocation is not computationally feasible. In Supplemental Appendix B, we show that it is not even possible to guarantee finding a stable allocation that achieves at least 94% of the size of the maximum-size stable allocation. Since our time-efficient algorithm is not optimal, the difference (in cardinality) between its result and the allocation that results from SR-DA is a lower bound on the maximal possible difference. Indeed, Appendix Table D1 presents the results of other stable algorithms which apply market power more broadly and achieve a larger increase in the number of assigned students in our empirical application.

## 5 Large Hungarian College Admissions Markets

We have shown that Hungarian college admissions markets may admit stable allocations where funding is not allocated based on merit, and that the rural hospital theorem and the consensus property do not extend to this domain. In this section, we ask what is required for these results to hold in large matching markets.

We find that when students' sensitivity to financial terms is heterogeneous (i.e., some students consider both contracts with the same college to be nearly perfect substitutes while others do not), the preference flip algorithm assigns more students to college relative to the outcome of DA (Theorem 3). By contrast, no stable allocation increases the number of students assigned to college when all students consider both contracts with the same college to be nearly perfect substitutes (Proposition 5) or when all students do not regard different contracts with the same college as close substitutes (Proposition 6).

### 5.1 A Model of Large Hungarian College Admissions Markets

**Random Hungarian college admissions markets** A *random Hungarian college admissions market* is a tuple  $\tilde{\Gamma} = \langle S, C, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C}, k, mode \rangle$ . It includes the sets of students and colleges, as well as colleges' preferences, rankings, and quotas. A random Hungarian college admissions market induces a Hungarian college admissions market by drawing students' preferences randomly. The randomization is governed by the integer  $k$ , which determines the number of contracts to be considered, and the parameter *mode*, which determines the mode of randomization.<sup>20</sup>

**Modes of randomization** We focus on two main modes of randomization, *independent* and *substitutes*. Under the *independent* mode of randomization, each student's preferences are randomly drawn independently and uniformly from the set of all permutations over  $2k$  contracts (other contracts are not acceptable). Under this mode of randomization, a typical student is extremely sensitive to financial terms (i.e., finds at most one set of financial terms acceptable in each college).

Under the *substitutes* mode of randomization, preferences are drawn as follows. First, for each student,  $k$  acceptable colleges are drawn uniformly at random. Second, a permutation over the  $2k$  contracts with these colleges is drawn independently and uniformly at random. This mode of randomization leads to substantial heterogeneity in students' sensitivity towards financial terms (i.e., some students rank the state-funded contract and the self-funded contract with the same college consecutively, while others do not).

We note that under the *substitutes* mode of randomization, there are no students who are extremely sensitive to financial terms (i.e., rank state-funded contract exclusively). The presence of such students would greatly simplify the proof of Theorem 3 and improve the bounds therein.<sup>21</sup>

<sup>20</sup>We build on the model of [Kojima and Pathak \(2009\)](#) that builds on the earlier work of [Immorlica and Mahdian \(2005, 2015\)](#) for large one-to-one matching markets (without contracts). Since we require colleges to have capacities larger than one we follow the notation and conditions of [Kojima and Pathak \(2009\)](#). [Hassidim et al. \(2017a\)](#) have shown that similar results hold for other large market models.

<sup>21</sup>[Rheingans-Yoo \(2020\)](#) proposes other ways to improve our bounds.

All of our results (for both modes of randomization) continue to hold when all (or most) students prefer state-funded seats to self-funded seats in the same program and when students' preferences are drawn from heterogeneous distributions (e.g., some students prefer economics majors while others prefer physics).

**Regular sequences of markets** A sequence of random Hungarian college admissions markets,  $\{\tilde{\Gamma}^n\}_{n=1}^{\infty}$  is *regular* if there exist integers  $k, \bar{q}$  and  $\lambda$ , all greater than one, and a mode of randomization *mode*, such that for each  $\Gamma^n = \langle S^n, C^n, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C^n}, k^n, mode^n \rangle$ :

1.  $|C^n| = n$  for all  $n$ ,
2.  $k^n = k$  and  $mode^n = mode$  for all  $n > 2k$ ,
3.  $0 < q_c^t \leq \bar{q}$  for all  $n, c \in C^n$  and  $t \in T$ ,
4.  $s \gg_c \emptyset$  for all  $n, c \in C^n$ , and  $s \in S^n$ ,
5.  $\frac{1}{\lambda}n \leq |S^n| \leq \lambda n$  for all  $n$ .

Condition 1 assures that the number of colleges grows as the sequence progresses. Condition 2 assures that the distribution of preferences is governed by the same mode of randomization and that the number of contracts that students consider acceptable is uniformly bounded on the sequence. Condition 3 assures that the number of positions of each type is positive and uniformly bounded across colleges and markets. Condition 4 assures that colleges find any student acceptable. Condition 5 assures that the number of students does not grow much faster or much slower than the number of colleges. [Kojima and Pathak \(2009\)](#) require only the first half of this condition. We require the second half as well since we are interested in instances where a substantial fraction of colleges have multiple stable allocations. In markets with a small number of students (who each find contracts with at most  $k$  colleges acceptable) most colleges will not even be a party to any individually rational allocation. The only other difference from the conditions of [Kojima and Pathak \(2009\)](#) is the requirement that colleges offer both types of seats (Condition 3). Had each college only offered one type of seat, both the *independent* and the *substitutes* modes of randomization would induce regular (uniform) [Kojima and Pathak](#) sequences of markets.<sup>22</sup> All of our results continue to hold if the fraction of colleges that offer multiple contracts is bounded away from zero.

## 5.2 Stable Allocations in Large Markets

**Theorem 3.** *For any regular sequence of markets with the substitutes mode of randomization, there exists  $\Delta > 0$ , such that comparing the stable allocation  $\tilde{Y}'_n$  to the outcome of SP-DA:*

1. *The expected number of assigned students is larger. Formally,*

$$\liminf_{n \rightarrow \infty} \frac{\mathbb{E} [|\tilde{Y}'_n|]}{\mathbb{E} [|\tilde{Y}_n^{SP-DA}|]} > 1.$$

<sup>22</sup>These sequences would also meet [Kojima and Pathak](#)'s thickness condition.

2. The expected fraction of newly assigned students, funding losers, and upward transfers are all bounded below by  $\Delta$ .

Finally, given access to students' preferences, and colleges' quotas and rankings of students,  $\tilde{Y}'_n$  can be computed time-efficiently (i.e., in polynomial-time) using the preference flip algorithm.

**Proposition 5.** *If all students in a Hungarian college admissions market are not sensitive to financial terms, then the same students are assigned across all stable allocations, and each college is assigned the same number of students across all stable allocations.*

We note that Proposition 5 is not probabilistic and does not rely on any of the large market assumptions.

**Proposition 6.** *Let  $\{\tilde{\Gamma}^n\}_{n=1}^\infty$  be a regular sequence of markets such that the mode of randomization is independent. Then, the expected fraction of colleges (students) that receive the same assignment in all stable allocations approaches 1 as the number of colleges,  $n$ , approaches infinity. Furthermore, with high probability, the fraction of colleges (students) that receive the same assignment in all stable allocations is arbitrarily close to 1 in sufficiently large markets.*

Proposition 6 states that the set of stable allocations is small when students do not regard different contracts with the same college as close substitutes. This result generalizes to a broad class of modes of randomization that preserve this property (see Supplemental Appendix B). These include modes of randomization where all (or most) students prefer state-funded seats to self-funded seats in the same program as well as modes where students' preferences are drawn from heterogeneous distributions (e.g., some students prefer economics majors while others prefer physics). In our empirical application, a large fraction of applicants rank the state-funded and self-funded contracts with the same college (see Section 6). This pattern is not consistent with any of these modes of randomization.

We conclude by noting that our model preserves the property of the Kojima and Pathak model that the outcomes of SP-DA and SR-DA are similar.

**Proposition 7.** *Let  $\{\tilde{\Gamma}^n\}_{n=1}^\infty$  be a regular sequence of markets such that the mode of randomization is independent or substitutes. Then, the expected fraction of colleges (students) that receive a different assignment under  $\tilde{Y}^{SP-DA}$  and  $\tilde{Y}^{SR-DA}$  approaches zero as the number of colleges,  $n$ , approaches infinity.*

## 6 Empirical Findings

In this section, we evaluate our theoretical predictions using data from Hungarian college admissions. We document heterogeneity in applicants' sensitivity to financial terms. Consistent with our theory, we find that the preference flip algorithm increases the number of applicants admitted to college relative to the outcome of SR-DA and that SP-DA and SR-DA are essentially identical.<sup>23</sup> Additionally, the preference flip algorithm benefits low socioeconomic status applicants and increases geographic mobility, especially to the periphery.

<sup>23</sup>Appendix G presents a similar analysis using data from the Israeli Psychology Master's Match, where an overwhelming majority of applicants are not sensitive to financial terms. The preference flip algorithm does not increase the number of admitted applicants.

## 6.1 Data

Our main data source is an administrative dataset that contains information on the bachelor's degree admissions process in Hungary. In particular, we observe each applicant's ROL, the priority score for each contract in the ROL, and all the information required to (re)calculate these priority scores.<sup>24</sup> This information includes grades in various subjects in the final two years of high school (11th and 12th grades), performance in the matriculation exams, and the number of points the applicant received for claiming a disadvantaged background. For each applicant we also observe gender, postal code, and a high-school identifier. For each alternative (program–terms pair) the data includes its realized priority-score cutoff (i.e., the minimal score that was required to gain admission to this alternative). We complement our data with hand-collected information on the program-specific capacities from the 2007 information booklet.

The second data source is the T-STAR dataset of the Hungarian Central Statistics Office. We use it to obtain settlement-level annual information on collected income taxes. In particular, we calculate the per-capita gross annual income for all 3,164 settlements in 2007 and merge it with our main administrative data based on settlement identifiers.

The third data source is the National Assessment of Basic Competencies (NABC). Following [Horn \(2013\)](#), we create an NABC-based SES index, which is a standardized measure that utilizes survey information of the NABC. The NABC-based SES index resembles the economic, social, and cultural status (ESCS) indicator of the OECD PISA survey. It combines three subindices: an index of parental education, an index of home possessions (number of bedrooms, mobile phones, cars, computers, books, etc.), and an index of parents' labor-market status. Since this data is available starting in 2008, we calculate the average value of students' NABC-based SES index for each high school between 2008 and 2012, and merge this with our main administrative dataset using high-school identifiers.<sup>25</sup>

Appendix Table [C2](#) summarizes the means and standard deviations of the background characteristics of applicants.

To set the stage for our main analysis, we first calculate the outcome of SR-DA. The SR-DA assignment replicates the realized assignment for all but 3.2 percent of the applicants (3,528 applicants).<sup>26</sup> We call the resulting allocation the benchmark. Appendix Table [C1](#)

---

<sup>24</sup>Our data report up to 7 contracts from each ROL: the first 6 contracts and the contract to which the applicant is assigned. The dataset also reports the number of contracts in each ROL. We observe the complete ROL for 91.7 percent of applicants and for 93.5 percent of all ranked contracts.

<sup>25</sup>The high-school-specific average NABC-based SES index is stable over time. Between 2008 and 2012, 95% of the variation in the high-school-specific averages of the NABC-based SES index is explained by high-school fixed effects. Furthermore, each year, about one-third of the student-level variation in the NABC-based SES index is explained by high-school fixed effects.

<sup>26</sup>There are small gaps and issues that do not allow us to fully replicate the allocation. First, we only observe up to 7 contracts in each ROL (corresponding to the top 6 choices and the realized allocation). Second, we do not know the results of auditions and other specialized admissions exams that are held by a small number of programs (e.g., art programs). Third, the mechanism has rules regarding cases where there are multiple (tied) marginal applicants which are not observable to us ([Biró, 2008](#); [Biró and Kiselgof, 2015](#)). We address this by breaking priority ties with a single lottery. Finally, the data contain some inconsistencies, for example, applicants with a priority score of zero assigned to selective programs. In our main analysis, we address these inconsistencies by fixing the assignment of these applicants to the realized assignment, and make the corresponding seats unavailable to others. Appendix [E](#) presents the results from other approaches,



shows that the benchmark allocation is larger by 2,567 assigned applicants, and that the difference is the result of increased utilization of self-funded seats. Intuitively, the higher utilization of self-funded seats reduces the scope for increasing the size of the stable allocation using the preference flip algorithm.

## 6.2 The Distribution of Applicants' Rank-Order Lists

Table 2 presents summary statistics on the characteristics of applicants' ROLs for all applicants, and for disadvantaged and non-disadvantaged applicants separately. Panel A presents statistics related to applicants' sensitivity towards funding. Panel B focuses on program characteristics, such as geographic location and field of study.

Panel A shows that 51.4 percent of applicants were extremely sensitive to financial terms: they ranked state-funded contracts exclusively. Other applicants were less sensitive to financial terms, including 11.6 percent of the applicants who ranked the same program with state-funding and self-funding consecutively on the top of their ROL.

Panel B shows that 49.4 percent of applicants rank exclusively contracts that are in the same settlement, 35 percent of applicants rank exclusively contracts at a single university, and 26.3 percent of applicants rank exclusively contracts at a single faculty of a university. Panel B also shows that 54.9 percent of applicants rank exclusively contracts in a single field of study, and 29.6 percent of applicants rank exclusively contracts in a single major.

Disadvantaged and non-disadvantaged applicants differ in terms of their sensitivity to financial terms and other program characteristics. Non-disadvantaged applicants were nearly twice as likely to rank the same program with state-funding and self-funding consecutively, and 29.4 percentage points less likely to rank state-funded contracts exclusively compared to disadvantaged applicants. Additionally, disadvantaged applicants were less likely to express preference for program characteristics.

Overall, these patterns suggest that disadvantaged and non-disadvantaged applicants trade off program characteristics (e.g., major or location) and financial terms differently. Disadvantaged applicants are more sensitive to financial terms, while non-disadvantaged applicants are more willing to pay for certain program characteristics. Since market power can only be exercised over applicants who rank two contracts in the same program, these patterns suggest that the direct effect of the preference flip algorithm would likely benefit disadvantaged applicants.

## 6.3 The Size of the Set of Stable Allocations

This section compares the stable allocations resulting from SR-DA, SP-DA, and the preference flip algorithm along several dimensions. We show that SR-DA and SP-DA yield essentially identical allocations, while the preference flip algorithm leads to meaningful changes: it changes the assignment of approximately 8 percent of college applicants benefiting low-SES applicants disproportionately, increases the number of applicants admitted to college by about 2 percent, and promotes geographic mobility relative to SR-DA and SP-DA.

---

and establishes the robustness of our findings.



Table 2: Characteristics of applicants' ROLs

	All applicants (%) (1)	Non-disadvantaged (%) (2)	Disadvantaged (%) (3)	p-value ((2)=(3)) (4)
<i>A. Preference for funding</i>				
State-funded contract exclusively	51.4	50.0	79.4	0.00
Self-funded contract exclusively	18.5	19.2	2.0	0.00
State- and self-funded contracts	30.1	30.8	18.6	0.00
Same study program consecutively	15.5	15.8	8.3	0.00
Same study program consecutively on the top of the ROL	11.6	11.9	6.2	0.00
<i>B. Preference for program characteristics</i>				
Single program location	49.4	50.0	36.6	0.00
Single university	35.0	35.4	27.4	0.00
Single faculty	26.3	26.9	15.3	0.00
Single field of study	54.9	55.6	40.1	0.00
Single major	29.6	30.3	15.5	0.00
# of applicants	108,854	103,840	5,014	

*Notes:* The table reports summary statistics on applicants' ROL. Panel A focuses on preference for funding. Specifically, Panel A shows the share of applicants who rank state-funded (self-funded) contracts exclusively, the share of students who both rank state-funded and self-funded contracts, and the share of applicants who rank the same study program with state-funding and self-funding consecutively, and who rank the same study program with state-funding and self-funding consecutively on the top of the ROL. Panel B focuses on preference for program characteristics. Specifically, Panel B shows the share of applicants who rank exclusively contracts that are in the same settlement (single program location), at a single university (single university), and at a single faculty of a university (single university), and the share of applicants who rank exclusively contracts in a single field of study, and in a single major. Column (1) presents these shares for all applicants, and columns (2) and (3) report these shares for non-disadvantaged and disadvantaged applicants, respectively. We test whether these shares differ between disadvantaged and non-disadvantaged applicants. Column (4) reports the corresponding p-values.

### 6.3.1 Winners, Losers, and the Number of Assigned Applicants

Table 3 compares the outcome of the preference flip algorithm to the SR-DA benchmark. Consistent with Theorem 3, Panel A shows that the preference flip algorithm increases the number of assigned applicants by 1.9 percent (1,558 applicants). The increase is concentrated in self-funded, full-time seats.

In the absence of the consensus property, it is interesting to understand how the preference flip algorithm affects different groups of students. Panel B provides statistics on the set of winners from the preference flip algorithm. We find that the preference flip algorithm generates 4,798 winners relative to SR-DA. Approximately 44 percent of the winners were not assigned under the benchmark (i.e., are newly assigned by the preference flip algorithm) and a similar number are assigned to a different program from that of their benchmark assignment. The rest of the applicants are assigned to the same program under preferable financial terms.

Panel C provides statistics on the set of losers from the preference flip algorithm. We find that the preference flip algorithm generates 4,260 losers relative to SR-DA. The majority

of losers (more than 70 percent) are assigned to the same program under less favorable financial terms. The rest of the losers are assigned to another program under a contract they ranked lower or become unassigned. Our theoretical analysis shows that the impact of the preference flip algorithm on redistribution is complex (Proposition 4). The direct effect of applying local market power is that some applicants lose their state funding and other applicants win by gaining admission to the program with state funding (instead of being unassigned or being assigned to a less desirable alternative). But, if the program already filled its capacity for self-funded applicants, this means that some applicants will lose their assignment to the unfunded contract. Additionally, displaced applicants (either winners or losers) can lead to chains of reassignments. This complexity is reflected in Panels B and C. Our empirical evaluation shows that, in spite of the theoretical possibility, a relatively small number of losers become unassigned or assigned to a new program.

We also evaluate what a single college can achieve by applying local market power. To this end, we focus on programs that can increase the size of their cohort (have unfilled self-funded capacity) and run a version of preference flip algorithm that allows only this program to apply local market power (by initializing  $A' = A \cap (S \times \{c\} \times T)$ ). We find that, on average, the number of applicants assigned to this program increases by 3.6. This number corresponds to a total increase of 1,741 applicants (compared to 1,558 when  $A' = A$ ).

Table 3: Winners, losers, and the number of assigned applicants

	Benchmark		Preference flip algorithm (3)
	SR-DA (1)	SP-DA (2)	
<i>A. Number of assigned applicants</i>			
Assigned to a contract	84,130	84,130	85,688
Assigned to a state-funded contract	48,725	48,725	48,709
Assigned to a self-funded contract	35,405	35,405	36,979
Assigned to a self-funded, full-time contract	13,321	13,321	14,418
<i>B. Winners</i>			
– Newly assigned		8	4,798
– New program		0	2,121
– Same program, preferred financial terms		8	2,282
		0	395
<i>C. Losers</i>			
– Newly unassigned		0	4,260
– New program		0	563
– Same program, less preferred financial terms		0	674
		0	3,023

*Notes:* The table presents the number of assigned applicants under various stable assignments (Panel A). Panels B and C describe the number of winners and losers from changing the benchmark SR-DA to SP-DA and to our alternate algorithms. An applicant is a winner if she is assigned to a contract she ranked higher than her contract in the benchmark assignment. An applicant is a loser if she is assigned to a contract she ranked lower than her assigned contract in the benchmark assignment or if she becomes unassigned. Column (1) shows the benchmark assignment. Column (2) presents the SP-DA assignment. Column (3) presents the preference flip algorithm assignment.

Column (2) of Table 3 presents the results of changing SR-DA to SP-DA. As predicted by Corollary 3, SP-DA and SR-DA assign the same number of applicants. Furthermore,

changing the assignment mechanism from SR-DA to SP-DA only changes the allocation of 8 applicants (approximately 0.01%). Consistent with Corollary 4, all 8 applicants prefer their SP-DA assignment.

According to Corollary 1, only the 8 applicants who strictly prefer the outcome of SP-DA to SR-DA have an incentive to misrepresent their preferences under SR-DA. This finding supports our reliance on reported preferences in our interpretation of winners and losers.<sup>27</sup>

### 6.3.2 Characteristics of Winners and Losers

Table 4 compares the characteristics of winners and losers for the preference flip algorithm. On average, winners' SES is lower than losers'. The share of disadvantaged applicants among winners is 6.0 percent relative to 2.5 percent among losers. Additionally, the average per-capita income in winners' settlements of residence is lower, the average NABC-based SES index in their high schools is higher, and they are less likely to live in the capital or a county capital.

Table 4 also reveals that, on average, winners have lower academic achievement than losers. The average 11th-grade GPA among winners is 3.5, relative to 3.7 among losers. The difference corresponds to 0.26  $(=(3.690-3.475)/0.84)$  standard deviations of the distribution of college applicants' 11th-grade GPA. In summary, applicants who prefer the assignment resulting from the preference flip algorithm to the merit-based benchmark are lower-achieving low-SES applicants.

### 6.3.3 Geographic Mobility

Table 5 studies geographic mobility. Geographic mobility is affected through winners who get newly assigned or get assigned to a new program at another location, and through the smaller number of losers who get unassigned or get assigned to a new program at another location. We compare the allocation resulting from the preference flip algorithm to the benchmark across three dimensions of geographic mobility. First, columns (1) and (2) focus on *movers*: applicants whose assigned study program is not located in the county where they reside. Second, columns (3) and (4) focus on applicants who get assigned to a study program in the periphery. Third, columns (5) and (6) focus on the subset of movers who move to the capital, that is, applicants who get assigned to a study program in the capital but do not reside there. Since low-SES applicants are more likely to reside outside the capital, and prestigious study programs are located in the capital, this latter dimension of geographic mobility may be interpreted as a measure of social mobility (Mari et al., 2020).

Columns (1) and (2) show that the preference flip algorithm increases the number of movers by 1,094 (from 51,221 to 52,315, see Panel A). This increase is largely explained by

---

<sup>27</sup>To verify that applicants had limited incentives to misrepresent their preferences, we also conducted a simulation study. We added to each applicants' ROL all contracts with 6 programs in fields that appear on the original ROL. We then drew a random ROL consisting of all these contracts uniformly at random and ran SP-DA and SR-DA. We repeated this process 100 times. On average, 7.1 applicants (out of the 108,854) could manipulate SR-DA. We repeated this analysis requiring that the funded contract in each program is ranked higher than the unfunded contract with the same program in all ROLs. On average, 7.4 applicants could manipulate SR-DA.

Table 4: Characteristics of winners and losers: preference flip algorithm

	Winners (1)	Losers (2)	p-values: (1)=(2) (3)
Disadvantaged	0.060	0.025	0.000
Per-capita annual gross income (1000 USD, 2007 prices)	9.719	10.399	0.000
NABC-based SES index	0.174	0.306	0.000
Capital	0.184	0.280	0.000
County capital	0.190	0.216	0.003
Town	0.330	0.293	0.000
Village	0.297	0.212	0.000
11th-grade GPA (1-5)	3.475	3.690	0.000
Female	0.571	0.552	0.076
Number of applicants	4,798	4,260	9,058

*Notes:* The table reports the mean values of characteristics of winners and losers from changing the benchmark SR-DA to the preference flip algorithm. An applicant is a winner if she is assigned to a contract she ranked higher than her benchmark assignment. An applicant is a loser if she is assigned to a contract she ranked lower than her assigned contract in the benchmark assignment or if she becomes unassigned. Column (3) presents p-values for the equality of mean characteristics of winners and losers.

the increase in the number of newly assigned applicants to a location outside their county of residence (1,330 newly assigned movers, see column (2) of Panel B). The difference is the result of 311 movers who become unassigned, an increase of 48 in the number of movers among winners who are assigned to a new program (=1,494–1,446), and an increase of 27 in the number of movers among losers who are assigned to a new program (=403–376).

Columns (3) and (4) show that the number of applicants assigned to the periphery increases by 1,110 (from 47,845 to 48,955, see Panel A). The increase in the number of applicants assigned to the periphery accounts for 71 percent of the increase in the number of assigned applicants. This finding is in line with the fact that most seats in the capital are filled in the benchmark assignment (see Appendix Table C8). Among winners, the number of applicants assigned to the periphery increases by 1,269 (=1,460+1,211–1,402), and, among losers, the number of applicants assigned to the periphery decreases by 159 (=339–233–265). Thus, unlike rural hospitals, peripheral colleges can improve the utilization of available seats.

Columns (5) and (6) show that the number of applicants who move to the capital increases by 381 (from 22,444 to 22,825, see Panel A). Thus, applicants from the periphery account for 85 percent (=381/448) of the increase in the number of applicants assigned to the capital. Among winners, the number of applicants who move to the capital increases by 590 (=435+694–539), and, among losers, the number of applicants who move to the capital decreases by 209 (=167–167–209).

## 7 Discussion

We study stable allocations in college admissions markets where students can attend the same college under different financial terms. We show that the deferred acceptance algorithm identifies a stable allocation where funding is allocated based on merit and that the set of *merit-based* stable allocations is small. By contrast, when students are heterogeneous

Table 5: Geographic mobility: Preference flip algorithm

	Mover		Assigned to periphery		Moved to capital	
	SR-DA (1)	PFA (2)	SR-DA (3)	PFA (4)	SR-DA (5)	PFA (6)
<i>A. Total</i>						
All assigned applicants	51,221	52,315	47,845	48,955	22,444	22,825
<i>B. Winners</i>						
Newly assigned (N= 2,121)	–	1,330	–	1,460	–	435
New program (N= 2,282)	1,446	1,494	1,402	1,211	539	694
<i>C. Losers</i>						
Newly unassigned (N= 563)	311	–	233	–	167	–
New program (N= 674)	376	403	265	339	209	167

*Notes:* The table presents the number of movers (i.e., applicants whose assigned study program is not located in the county where they reside), of applicants assigned to the periphery (i.e., not to the capital), and of applicants who moved to the capital (i.e., applicants who get assigned to a study program in the capital but do not reside there) under the preference flip algorithm (PFA) and in the benchmark SR-DA assignments. Panel A presents the total number of applicants with these characteristics. Panels B and C focus on winners and losers who are assigned to a different study program.

in the way they trade off program characteristics and contractual terms, the set of stable allocations is large and different stable allocations differ in the number of students assigned to college.

While our focus was on financial aid, our results generalize to setting with other contractual terms. In the context of college admissions, such terms include specialized study tracks (e.g., a business school offering management and accounting tracks) or access to dormitories. Beyond college admissions, financial terms may correspond to different service terms in the military (Sönmez, 2013) or different levels of compensation (Niederle, 2007).

Our findings raise several important questions. First, *how does one implement alternative stable allocations?* This is challenging, because non-merit-based stable mechanisms give academically strong applicants an incentive to overstate their sensitivity to financial terms. One approach for addressing this challenge is to use hard information. For example, governments may use existing administrative data or facilitate the collection of information by colleges through forms like the Free Application for Federal Student Aid (FAFSA). American college applicants who complete this form that determines their eligibility for federal financial aid can choose to forward this information to particular colleges. Similarly, enrollment management and personalized financial aid packages are widespread in the U.S. college admissions market, and private companies are offering financial aid optimization tools using large proprietary datasets (Lieber, 2021).

Another interesting question is *how to compute a stable allocation that supports the objectives of a policymaker.* Policymakers may have different objectives (e.g., maximize the number of students assigned, diversity, etc.). We have shown that—at least for the objective of maximizing the number of assigned students—this computation may be hard. Still, in college admissions markets, a long waiting period until the release of official result is common, and this may allow solving realistic-size problems in practice. Leveraging our characterization of



stability (Proposition 1), integer programming methods in the spirit of [Ágoston et al. \(2016\)](#) and [Delorme et al. \(2019\)](#) are a promising direction.

## References

- Abdulkadiroğlu, Atila, and Tayfun Sönmez.** 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729–747.
- Abizada, Azar.** 2016. "Stability and Incentives for College Admissions with Budget Constraints." *Theoretical Economics*, 11(2): 735–756.
- Afacan, Mustafa Oğuz.** 2020. "Graduate Admission with Financial Support." *Journal of Mathematical Economics*, 87 114–127.
- Ágoston, Kolos Csaba, Péter Biró, and Iain McBride.** 2016. "Integer Programming Methods for Special College Admissions Problems." *Journal of Combinatorial Optimization*, 32(4): 1371–1399.
- Ashlagi, Itai, Mark Braverman, and Avinatan Hassidim.** 2014. "Stability in Large Matching Markets with Complementarities." *Operations Research*, 62(4): 713–732.
- 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.
- Ashlagi, Itai, and Afshin Nikzad.** 2020. "What Matters in School Choice Tie-Breaking? How Competition Guides Design." *Journal of Economic Theory*, 190, p. 105120.
- Ashlagi, Itai, Afshin Nikzad, and Assaf Romm.** 2019. "Assigning More Students to Their Top Choices: A Comparison of Tie-Breaking Rules." *Games and Economic Behavior*, 115 167–187.
- Aygün, Orhan, and Bertan Turhan.** 2017. "Large-Scale Affirmative Action in School Choice: Admissions to IITs in India." *American Economic Review*, 107(5): 210–13.
- Aygün, Orhan, and Bertan Turhan.** 2020. "Dynamic Reserves in Matching Markets." *Journal of Economic Theory*, 188, 105069.
- Aygün, Orhan, and Inácio Bó.** 2016. "College Admission with Multidimensional Privileges: The Brazilian Affirmative Action Case." Mimeo.
- Azevedo, Eduardo M., and Eric Budish.** 2018. "Strategy-Proofness in the Large." *The Review of Economic Studies*, 86(1): 81–116.
- 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.
- 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, Tamás Fleiner, Robert W. Irving, and David F. Manlove.** 2010. "The College Admissions Problem with Lower and Common Quotas." *Theoretical Computer Science*, 411(34): 3136–3153.
- Biró, Péter.** 2008. "Student Admissions in Hungary as Gale and Shapley Envisaged." Technical Report TR-2008-291, University of Glasgow.
- Biró, Péter, and Sofya Kiselgof.** 2015. "College Admissions with Stable Score-Limits." *Central European Journal of Operations Research*, 23(4): 727–741.
- Caniglia, Alan S., and Daniel R. Porterfield.** forthcoming. "Addressing Under-Matching in



- College Enrollment: Toward an Economic Rationale and a Case Study." In *More Equal by Design: Economic Design Responses to Inequality* (S. D. Kominers and A. Teytelboym, editors). OUP.
- Coles, Peter, and Ran I. Shorrer.** 2014. "Optimal Truncation in Matching Markets." *Games and Economic Behavior*, 87: 591–615.
- Delacrétaz, David, Scott Duke Kominers, and Alexander Teytelboym.** 2016. "Refugee Resettlement." *University of Oxford Department of Economics Working Paper*.
- Delorme, Maxence, Sergio García, Jacek Gondzio, Joerg Kalcsics, David Manlove, and William Pettersson.** 2019. "Mathematical Models for Stable Matching Problems with Ties and Incomplete Lists." *European Journal of Operational Research*, 277(2): 426–441.
- Demange, Gabrielle, David Gale, and Marilda Sotomayor.** 1987. "A Further Note on the Stable Matching Problem." *Discrete Applied Mathematics*, 16(3): 217–222.
- DePalma, Anthony.** 1992. "M.I.T. Ruled Guilty in Antitrust Case." Sep 3, URL: [www.nytimes.com/1992/09/03/us/mit-ruled-guilty-in-antitrust-case.html](http://www.nytimes.com/1992/09/03/us/mit-ruled-guilty-in-antitrust-case.html).
- Dubins, Lester E., and David A. Freedman.** 1981. "Machiavelli and the Gale–Shapley Algorithm." *American Mathematical Monthly*, 88(7): 485–494.
- Dur, Umut, Scott D. Kominers, Parag A. Pathak, and Tayfun Sönmez.** 2018. "Reserve Design: Unintended Consequences and the Demise of Boston's Walk Zones." *Journal of Political Economy*, 126(6): 2457–2479.
- 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.
- Ágoston, Kolos Csaba, Péter Biró, Endre Kováts, and Zsuzsanna Jank.** 2021. "College admissions with ties and common quotas: Integer programming approach." *European Journal of Operational Research*.
- Halldórsson, Magnús M., Kazuo Iwama, Shuichi Miyazaki, and Hiroki Yanagisawa.** 2007. "Improved Approximation Results for the Stable Marriage Problem." *ACM Transactions on Algorithms (TALG)*, 3(3): , p. 30.
- Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer.** 2017a. "Need vs. Merit: The Large Core of College Admissions Markets." Mimeo.
- Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer.** 2017b. "Redesigning the Israeli Psychology Master's Match." *American Economic Review*, 107(5): 205–209.
- Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer.** 2019. "Contracts Are Not Salaries in the Hidden-Substitutes Domain." *Economics Letters*, 181: 40–42.
- 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 Scott D. Kominers.** 2015. "Hidden Substitutes." In *Proceedings of the 16th ACM Conference on Economics and Computation.*, p. 37.
- Hatfield, John W., and Paul R. Milgrom.** 2005. "Matching with Contracts." *American Eco-*

- nomic Review*, 95(4): 913–935.
- Hatfield, John William, Scott Duke Kominers, and Alexander Westkamp.** 2020. “Stability, Strategy-Proofness, and Cumulative Offer Mechanisms.” *The Review of Economic Studies*, 88(3): 1457–1502.
- Heo, Eun Jeong.** 2017. “Financial Aid in College Admissions: Need-Based versus Merit-Based.” Mimeo.
- 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, No. 7754.
- Immorlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier.** 2020. “Information Acquisition in Matching Markets: The Role of Price Discovery.” Mimeo.
- Immorlica, Nicole, and Mohammad Mahdian.** 2005. “Marriage, Honesty, and Stability.” In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. 53–62, Society for Industrial and Applied Mathematics.
- Immorlica, Nicole, and Mohammad Mahdian.** 2015. “Incentives in Large Random Two-Sided Markets.” *ACM Transactions on Economics and Computation*, 3(3): 14–14.
- Impagliazzo, Russell, and Valentine Kabanets.** 2010. “Constructive Proofs of Concentration Bounds.” In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques.*: Springer, 617–631.
- Kelso, Alexander S., and Vincent P. Crawford.** 1982. “Job Matching, Coalition Formation, and Gross Substitutes.” *Econometrica*, 50(6): 1483–1504.
- Kim, Matthew.** 2010. “Early Decision and Financial Aid Competition among Need-Blind Colleges and Universities.” *Journal of Public Economics*, 94(5–6): 410–420.
- 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): 608–627.
- Kojima, Fuhito, Parag A. Pathak, and Alvin E. Roth.** 2013. “Matching with Couples: Stability and Incentives in Large Markets.” *The Quarterly Journal of Economics*, 128(4): 1585–1632.
- Konishi, Hideo, and M. Utku Ünver.** 2006. “Games of Capacity Manipulation in Hospital-intern Markets.” *Social Choice & Welfare*, 27(1): 3–24.
- Lee, SangMok.** 2017. “Incentive Compatibility of Large Centralized Matching Markets.” *The Review of Economic Studies*, 84(1): 444–463.
- Leshno, Jacob D, and Irene Lo.** 2020. “The Cutoff Structure of Top Trading Cycles in School Choice.” *The Review of Economic Studies*.
- Leyton-Brown, Kevin, Paul Milgrom, and Ilya Segal.** 2017. “Economics and Computer Science of a Radio Spectrum Reallocation.” *Proceedings of the National Academy of Sciences*, 114(28): 7202–7209.
- Lieber, Ron.** 2021. *The Price You Pay for College: An Entirely New Road Map for the Biggest Financial Decision Your Family Will Ever Make.*: Harper.

- Manlove, David F., Robert W. Irving, Kazuo Iwama, Suichi Miyazaki, and Yasufumi Morita.** 2002. "Hard Variants of Stable Marriage." *Theoretical Computer Science*, 276(1–2): 261–279.
- Mari, Tanaka, Narita Yusuke, and Moriguchi Chiaki.** 2020. "Meritocracy and Its Discontent: Long-run Effects of Repeated School Admission Reforms." Discussion papers 20002, Research Institute of Economy, Trade and Industry (RIETI).
- Nei, Stephen, and Bobak Pakzad-Hurson.** 2016. "Strategic Disaggregation in Matching Markets." Mimeo.
- 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." Mimeo.
- Pathak, Parag A., Alex Rees-Jones, and Tayfun Sönmez.** 2020. "Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms." National Bureau of Economic Research, No. 26767.
- Rheingans-Yoo, Ross.** 2020. "Large Random Matching Markets with Localized Preference Structures Can Exhibit Large Cores." Mimeo.
- Romm, Assaf, Alvin E. Roth, and Ran I. Shorrer.** 2020. "Stability vs. No Justified Envy." Mimeo.
- Roth, Alvin E.** 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research*, 7(4): 617–628.
- Roth, Alvin E.** 1984a. "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.** 1984b. "Stability and Polarization of Interests in Job Matching." *Econometrica*, 52(1): 47–57.
- Roth, Alvin E.** 1985. "The College Admissions Problem is not Equivalent to the Marriage Problem." *Journal of Economic Theory*, 36(2): 277–288.
- 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 Design." *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 Marilda A. Oliveira Sotomayor.** 1990. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society monographs, New York: Cambridge University Press.
- Roth, Alvin E., and Xiaolin Xing.** 1994. "Jumping the Gun: Imperfections and Institutions Related to the Timing of Market Transactions." *American Economic Review*, 84(4): 992–1044.
- Schwarz, Michael, and M. Bumin Yenmez.** 2011. "Median Stable Matching for Markets with Wages." *Journal of Economic Theory*, 146(2): 619–637.
- Shorrer, I. Ran, and Sándor Sóvágó.** 2018. "Obvious Mistakes in a Strategically Simple Col-

- lege Admissions Environment: Causes and Consequences." Mimeo.
- Sönmez, Tayfun.** 1997. "Manipulation via Capacities in Two-Sided Matching Markets." *Journal of Economic Theory*, 77(1): 197–204.
- Sönmez, Tayfun.** 1999. "Can Pre-arranged Matches Be Avoided in Two-Sided Matching Markets?" *Journal of Economic Theory*, 86(1): 148–156.
- Sönmez, Tayfun.** 2013. "Bidding for Army Career Specialties: Improving the ROTC Branching Mechanism." *Journal of Political Economy*, 121(1): 186–219.
- Storms, Evan.** 2013. "Incentives and Manipulation in Large Market Matching with Substitutes." Mimeo.
- Westkamp, Alexander.** 2013. "An Analysis of the German University Admissions System." *Economic Theory*, 53(3): 561–589.
- Yenmez, M. Bumin.** 2018. "A College Admissions Clearinghouse." *Journal of Economic Theory*, 176: 859–885.

## Appendix A Proofs

*Proof of Proposition 1.* If  $Y$  violates one of Conditions 1–4, then it is not stable. For the other direction, assume  $Y$  is not stable. We show that if  $Y$  violates none of Conditions 1, 2, or 4, then it must violate Condition 3.

Since  $Y$  is individually rational (by Condition 1), the allocation is blocked through some  $Z \neq \emptyset$ . By Condition 4,  $Z'$  contains a contract  $(\bar{s}, c, \bar{t})$  such that  $(\bar{s}, c, 1 - \bar{t}) \notin Y$ . Furthermore,  $Y$  is also blocked through  $Z' = \text{Ch}_c(Y \cup Z) \setminus Y$ , which contains  $(\bar{s}, c, \bar{t})$ .

Let  $(s, c, t) \in Z'$  be a contract with the highest ranked student according to  $\gg_c$  (formally,  $\hat{s} \gg_c s \implies (\{\hat{s}\} \times \{c\} \times T) \cap Z' = \emptyset$ ). By Condition 2,  $Y$  is not blocked through  $\{(s, c, t)\}$ . Therefore, the allocation  $Y$  assigns  $q_c^t$  students to  $c$  under the financial terms  $t$ , and all of them are ranked higher than  $s$  according to  $\gg_c$ . Therefore, since  $(s, c, t) \in \text{Ch}(Y \cup Z')$ , there must exist a contract  $(s', c, 1 - t) \in Z'$  such that  $(s', c, t) \in Y$ . We claim that the allocation must be blocked through  $\{(s, c, t), (s', c, 1 - t)\}$  (in violation of Condition 3).

To see this, note that if  $|Y \cap (S \times \{c\} \times \{1 - t\})| < q_c^{1-t}$  (i.e.,  $c$ 's  $(1 - t)$ -quota is not full), then  $c$  would prefer to add the contract  $(s, c, 1 - t)$  ( $s$  is acceptable to  $c$  since all contracts in  $Z'$  are), and swaps in the identities of the recipients of financial aid are less important to the college than the composition of the cohort. Otherwise, both of  $c$ 's quotas are full. Thus, there exists some  $s''$  such that  $(s'', c, 1 - t) \in Y$  and  $s'' \notin [\text{Ch}_c(Y \cup Z')]_S$ . This holds since  $s$  (a new student) must displace someone (both quotas are full), and all students who get assigned to  $c$  under financial terms  $t$  in  $Y$  are ranked higher than  $s$  according to  $\gg_c$  (hence, for  $(s, c, t)$  to be chosen by  $c$ , a contract with each of them must also be chosen). We denote by  $s^*$  be the lowest ranked student according to  $\gg_c$  such that  $(s^*, c, 1 - t) \in Y$  and  $s^* \notin [\text{Ch}_c(Y \cup Z')]_S$ . Finally, note that  $s' \neq s^*$ , and by the properties of  $\succ_c$  we have  $\text{Ch}_c(Y \cup \{(s, c, t), (s', c, 1 - t)\}) = ((Y \cap (S \times \{c\} \times T)) \cup \{(s, c, t), (s', c, 1 - t)\}) \setminus \{(s', c, t), (s^*, c, 1 - t)\}$ .  $\square$

*Proof of Proposition 2.* On the run of SP-DA, colleges never face a choice from a set that includes two contracts with the same student. Consequently, colleges' choices on the run of the algorithm coincide with the union of the choices of the two corresponding auxiliary colleges in the related market. Hence, the corresponding matching of  $Y^{\text{SP-DA}}$  in the related market is the student-optimal stable matching.  $\square$

*Proof of Corollary 1.* A student who benefits from misrepresenting her preferences to SP-DA (SR-DA) can do the same in the related market (by Proposition 2). The corollary, therefore, follows from the results of Demange et al. (1987), Dubins and Freedman (1981), and Roth (1982) for the related market.  $\square$

*Proof of Theorem 1. If  $Y$  corresponds to a stable matching then  $Y$  is certainly stable:* We prove the counterpositive: that if  $Y$  is not certainly stable, then the corresponding matching is not stable. Let  $Y$  be a feasible allocation. If  $Y$  is not certainly stable, then it must violate at least one of Conditions 1–3 of Proposition 1 or Condition 4'. If  $Y$  violates Condition 1 ( $Y$  is not individually rational), then either some student is assigned an unacceptable contract, or some college violates its quotas, or some college is assigned an unacceptable student. In all cases the corresponding matching also violates individual rationality. If  $Y$  violates Condition 2 ( $Y$  is blocked through  $\{(s, c, t)\}$ ), then either  $|Y \cap (S \times \{c\} \times \{t\})| < q_c^t$  or there exists



$(s', c, t) \in Y$  such that  $s \gg_c s'$ . In either case,  $s$  and  $(c, t)$  block the corresponding matching in the related market. If  $Y$  violates Condition 3 ( $Y$  is blocked through  $\{(s', c, 1 - t), (s, c, t)\}$  such that  $(s, c, 1 - t) \notin Y$  and  $(s', c, t) \in Y$ ), then the student  $s'$  and the auxiliary college  $(c, 1 - t)$  block the corresponding matching. Finally, if an individually rational allocation,  $Y$ , violates Condition 4' then either there exists a college,  $c$  and a pair of students that want to swap their terms in this college (in which case the higher ranked student according to  $\gg_c$  and the auxiliary college that this student prefers form a blocking pair in the related market), or there exists a college  $c$  where some student prefers to get different terms, and the quota of these terms is not full (in which case, the student and the auxiliary college that she prefers form a blocking pair in the related market).

**If  $Y$  corresponds to a stable matching in the related market,  $Y$  allocates funding based on merit:** Let  $Y$  be a stable allocation that corresponds to a stable matching in the related two-sided matching market. Then, for each  $c$  and  $t$ , if  $s$  is matched with  $(c, t)$  and  $s' \gg_c s$  then  $s'$  must weakly prefer her assignment to  $(c, t)$ .

**If  $Y$  is merit-based certainly stable, then it corresponds to a stable matching in the related market:** Since  $Y$  is stable, the corresponding matching cannot violate individual rationality. Similarly, if it is blocked by  $s$  and  $(c, t)$  it must be that  $(s, c, 1 - t) \in Y$ , since otherwise  $Y$  would be blocked by  $(s, c, t)$ . However, by Condition 4' if  $(s, c, 1 - t) \in Y$  and  $(s, c, t) \succ_s (s, c, 1 - t) \in Y$  then  $q_c^t$  students are assigned to  $c$  under  $Y$  and since  $Y$  is merit-based, all of these students are ranked higher than  $s$  according to  $\gg_c$ .

**If  $Y$  is merit-based stable, then there exists a merit-based certainly stable allocation that all students weakly prefer:** If  $Y$  is certainly stable, we are done. Otherwise, Condition 4' is violated. Since funding is allocated based on merit, this means that there exist  $s \in S$  and  $c \in C$  such that  $(s, c, t) \in Y$ ,  $(s, c, 1 - t) \succ_s (s, c, t)$ , and  $Y \cap (S \times \{c\} \times 1 - t) < q_c^{1-t}$  (but the college prefers that  $s$  is assigned under  $t$ ). Let  $s'$  be the highest ranked such student according to  $\gg_c$ . Let  $Y'$  denote the allocation  $Y \setminus \{(s', c, t)\} \cup \{(s', c, 1 - t)\}$ . This allocation is stable as otherwise  $Y$  is not stable (it is blocked through the same set of contracts, or through the same set of contracts together with  $\{(s', c, 1 - t)\}$ ). It is merit-based since  $s'$  is maximal with respect to  $\gg_c$ . If  $Y'$  is certainly stable, we are done. Otherwise, repeat the process. Since in each iteration all students are weakly better off and one student is (strictly) better off, the process will reach a certainly stable allocation after a finite number of steps (Condition 4' cannot be violated if all students receive their first choices). □

*Proof of Corollary 2.* Follows from [Gale and Shapley \(1962\)](#) by Theorem 1 and Proposition 2. □

*Proof of Corollary 3.* Follows from the rural hospital theorem for the related market. □

*Proof of Corollary 4.* Follows from the consensus property in the related market. □

*Proof of Proposition 3.* Follows from Example 1 in the main text and Example 2 below. □

**Example 2.** There are three students,  $S = \{r, p, g\}$ , and two colleges,  $C = \{h, c\}$ . College  $h$  has two seats, but only one of these seats is state-funded ( $q_h^0 = 1, q_h^1 = 1$ ), and college  $c$  has a single state-funded seat ( $q_c^0 = 0, q_c^1 = 1$ ). Both colleges rank  $r$  first,  $p$  second, and  $g$  third

(i.e.,  $r \gg_h p \gg_h g$  and  $r \gg_c p \gg_c g$ ). Students' preferences are  $(r, h, 1) \succ_r (r, h, 0) \succ_r \emptyset$ ,  $(g, h, 1) \succ_g (g, h, 0) \succ_g \emptyset$ , and  $(p, h, 1) \succ_p (p, c, 1) \succ_p \emptyset$ .

Under  $Y^{\text{SP-DA}} = \{(r, h, 1), (p, c, 1), (g, h, 0)\}$ ,  $r$  and  $g$  are assigned to  $h$ , and  $p$  is assigned to  $c$ . An alternative stable allocation,  $\{(r, h, 0), (p, h, 1)\}$ , has the two highest-ranked students,  $r$  and  $p$ , assigned to  $h$ , while the lowest-ranked student,  $g$ , remains unassigned.

*Proof of Proposition 4.* The algorithm terminates since the finite set  $A'$  becomes smaller in each iteration. If it stops while  $A'$  is not empty, the output is stable by construction. Otherwise, it is stable as it coincides with SP-DA. To see that the resulting stable allocation may not allocate funding based on merit note that it selects the non-merit-based stable allocation in Example 1.  $\square$

*Proof of Theorem 2.* We prove that Theorem 2 holds even if colleges are restricted to offer no more than one seat under each of the financial terms. Our proof relies on a reduction—we show that an algorithm that identifies a stable (certainly stable) allocation of maximum size quickly (in running time polynomial in the size of the input) can also be used to quickly produce a solution to another problem that is known to be NP-hard. Therefore, identifying such an algorithm proves that  $P = NP$ . The construction is provided in Supplemental Appendix B.  $\square$

### Proof of Theorem 3

**Definition 1.** For  $k + 1$  distinct colleges,  $c, c'_1, c'_2, \dots, c'_k \in C^n$ , and  $q_c^1 + q_{c'_1}^1 + q_{c'_1}^0 + 2(k - 1)\bar{q} + 1$  distinct students,  $\{r_i\}_{i=1}^{q_c^1}$ ,  $p$ ,  $\{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}$ ,  $w$ , and  $\{\{s_i^j\}_{i=1}^{2\bar{q}}\}_{j=2}^k$  let the event

$$E^n \left( c, \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \{\{s_i^j\}_{i=1}^{2\bar{q}}\}_{j=2}^k \right)$$

denote the case where:

1.  $r_1 \gg_c r_2 \gg_c \dots \gg_c r_{q_c^1} \gg_c p$
2.  $s_1^1 \gg_{c'_1} s_2^1 \gg_{c'_1} \dots \gg_{c'_1} s_{q_{c'_1}^0 + q_{c'_1}^1 - 1}^1 \gg_{c'_1} w$  and  $p \gg_{c'_1} w$
3.  $s_1^j \gg_{c'_j} s_2^j \gg_{c'_j} \dots \gg_{c'_j} s_{2\bar{q}}^j \gg_{c'_j} w$  for each  $2 \leq j \leq k$
4. The only students who find contracts with  $c$  acceptable are  $\{r_i\}_{i=1}^{q_c^1}$  and  $p$ . Formally, for all  $s \in S^n \setminus \left( \{r_i\}_{i=1}^{q_c^1} \cup \{p\} \right)$  and all  $t \in T$ ,  $\emptyset \succ_s (s, c, t)$ .
5. The only students who find contracts with  $c'_1$  acceptable are  $p$ ,  $\{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}$ , and  $w$ . Formally,  $\emptyset \succ_s (s, c'_1, t)$  for all  $s \in S^n \setminus \left( \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1} \cup \{p, w\} \right)$  and all  $t \in T$ .

6. For  $j \geq 2$ , the only students who find contracts with  $c'_j$  acceptable are  $\{s'_i\}_{i=1}^{2\bar{q}}$  and  $w$ . Formally,  $\emptyset \succ_s (s, c'_j, t)$  for all  $s \in S^n \setminus (\{s'_i\}_{i=1}^{2\bar{q}} \cup \{w\})$  and all  $t \in T$ .
7. Each  $r_j$  prefers  $(r_j, c, 1)$  to  $(r_j, c, 0)$  and prefers  $(r_j, c, 0)$  to any other contract. Formally,  $(r_j, c, 1) \succ_{r_j} (r_j, c, 0) \succ_{r_j} (r_j, u, t)$  for each  $r_j \in \{r_i\}_{i=1}^{q_c^1}$ , and  $u \in C^n \setminus \{c\}$  and  $t \in T$ .
8.  $p$  prefers  $(p, c, 1)$  to  $(p, c'_1, 1)$  and prefers  $(p, c'_1, 1)$  to any other contract. Formally,  $(p, c, 1) \succ_p (p, c'_1, 1) \succ_p z$  for each  $z \in (\{p\} \times C^n \times T) \setminus \{(p, c, 1), (p, c'_1, 1)\}$ .
9.  $(w, c'_1, 1) \succ_w (w, c'_2, 1) \succ_w \dots \succ_w (w, c'_k, 1) \succ_w \emptyset$ .
10. Each  $s'_i$  prefers  $(s'_i, c'_j, 1)$  to  $(s'_i, c'_j, 0)$  and prefers  $(s'_i, c'_j, 0)$  to any other contract. Formally,  $(s'_i, c'_j, 1) \succ_{s'_i} (s'_i, c'_j, 0) \succ_{s'_i} (s'_i, u, t)$  for each  $2 \leq j \leq k$ ,  $s'_i \in \{s'_i\}_{i=1}^{\bar{q}}$ , and  $u \in C^n \setminus \{c\}$  and  $t \in T$ .

**Lemma 1.** *There exists  $F > 0$  so that, so long as Conditions 1–3 of Definition 1 are met, the event  $E^n \left( c, \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=1}^{q_c^1}, \{s'_i\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \{\{s'_i\}_{i=1}^{2\bar{q}}\}_{j=2}^k \right)$  is realized with probability greater than  $F/n^{q_c^1 + q_{c'_1}^1 + q_{c'_1}^0 + 2(k-1)\bar{q} + k + 1}$ .*

*Proof.* First, observe that Conditions 1–3 depend on colleges rankings exclusively, so students' preferences are independent from them. Second, it is well-known since [Immorlica and Mahdian \(2005\)](#) that conditions of the form “ $m$  schools are not ranked by some subset of students” have a probability that is bounded below by a constant when each student is interested in a small number of colleges.<sup>28</sup> Third, the unconditional probability that a student's preferences meet Condition 7 or 10 is greater than  $f/n$  for  $f > 0$  (that depends on  $k$ ), and the probability conditional of Conditions 4–6 being met is slightly greater. Similarly, the unconditional probability for the preferences of  $p$  (respectively  $w$ ) to meet Condition 8 (respectively 9) is greater than  $f/n^2$  (respectively  $f/n^k$ ) and the probability conditional of Conditions 4–6 being met is slightly greater. Since students' preferences are drawn independently, the product of the lower bounds is a valid lower bound for the probability of the event.  $\square$

**Lemma 2.** *Given a Hungarian college admissions market, with  $s \gg_c \emptyset$  for all  $s \in S$  and  $c \in C$ , there are at least  $|S| \cdot (|S| - 1)/4$  pairs of distinct students such that both are ranked outside the top decile of  $\gg$  by at least 40 percent of the colleges.*

*Proof.* We prove this combinatorial claim using the probabilistic method. First, note that the fraction of pairs of students who both appear in the bottom 9 deciles of  $\gg$  of at least  $4n/10$  colleges is equal to the probability that this condition is satisfied by a random pair of students drawn uniformly at random from all possible distinct pairs. Let  $\mathbb{1}_c(s, s')$  equal 1

<sup>28</sup>The probability that  $m$  schools are not ranked by any student is greater than  $(1 - mk/n)^{\lambda n}$  converges to  $e^{-\lambda mk}$ .

if  $s$  and  $s'$  are in the top 10 percent of  $\gg_c$  and zero otherwise. Then, if a pair of distinct students  $(\mathbf{s}, \mathbf{s}')$  is drawn uniformly at random, the expected value of  $\mathbb{1}_c(\mathbf{s}, \mathbf{s}')$  is lower than 0.2 using (using Bonferroni's inequalities). By Markov's inequality

$$\Pr \left[ \sum_{c \in C} \mathbb{1}_c(\mathbf{s}, \mathbf{s}') > 0.4 |C| \right] \leq \frac{0.2 |C|}{0.4 |C|} = \frac{1}{2}$$

which completes the proof.  $\square$

**Lemma 3.** *There exists  $M > 0$  such that there are more than  $M \cdot n^{q_c^1 + q_{c_1}^1 + q_{c_1}^0 + 2(k-1)\bar{q} + k + 2}$  ways to select  $\left( c, \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c_1}^1 + q_{c_1}^0 - 1}, \left\{ \{s_i^j\}_{i=1}^{2\bar{q}} \right\}_{j=2}^k \right)$  such that Conditions 1–3 of Definition 1 are met.*

*Proof.* By Lemma 2 there are at least  $(n/\lambda - 2)^2/4$  ways to choose pairs of students that are ranked outside the top decile by at least  $n/4$  colleges. Given such a pair of students, let  $c$  and  $\{c'_j\}_{j=1}^k$  be colleges such that both students are ranked at the bottom 90 percent (three are at least  $\alpha \cdot n^{k+1}$  ordered ways the choose them, for  $\alpha > 0$ ). Label the member of the pair that is ranked lower in  $c'_1$  as  $w$  and the other student as  $p$ . Choose  $\{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c_1}^1 + q_{c_1}^0 - 1}$ , and  $\left\{ \{s_i^j\}_{i=1}^{2\bar{q}} \right\}_{j=2}^k$  distinct students who appear at the top 10 percent of the corresponding colleges ( $c$  for  $r_i$ 's and  $c'_j$  for  $s_i^j$ 's). There at least  $\beta \cdot n^{q_c^1 + q_{c_1}^1 + q_{c_1}^0 - 1 + 2(k-1)\bar{q}}$  possible selections, for some  $\beta > 0$ . Multiplying the three bounds we get the required result.  $\square$

**Lemma 4.** *For any college  $c$ , and any two different selections,  $z$  and  $z'$ , of*

$$\left( \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c_1}^1 + q_{c_1}^0 - 1}, \left\{ \{s_i^j\}_{i=1}^{2\bar{q}} \right\}_{j=2}^k \right),$$

*the events  $E^n(c, z)$  and  $E^n(c, z')$  are disjoint. Furthermore, there exists a polynomial-time algorithm that identifies the set  $\{z \mid E^n(c, z) \text{ is realized}\}$ .*

*Proof.* We describe a polynomial-time algorithm and show that it must identify the only realized event of the form  $E^n(c, z)$ , or determine that no such event is realized.

### Algorithm

1. Find a student,  $s$ , that ranks  $(s, c, 1)$  first and  $(s, c', 1)$  second, for  $c' \neq c$ . If no such student exists, declare that the set is empty and stop. Otherwise, this student is the candidate  $p$ , and the college  $c'$  is the candidate  $c'_1$ .
2. Find all students other than the candidate  $p$  that rank contracts with  $c$ . If their number differs from  $q_c^1 + 1$  declare that the set is empty and stop. Otherwise, the lowest ranked one according to  $\gg_c$  is the candidate  $w$ . If one of the other  $q_c^1$  students does not rank the funded contract with  $c$  first and the unfunded contract with  $c$  second, declare that

the set is empty and stop. Otherwise, they are the candidate  $\{r_i\}_{i=1}^{q_c^1}$  (with indices determined based on  $\gg_c$  and Condition 1 of Definition 1).

3. If the candidate  $w$  ranks some funded contract higher than the funded contract with the candidate  $c'_1$ , declare that the set is empty and stop. Otherwise, set the candidate  $\{c'_j\}_{j=2}^k$  as the remaining colleges (with indices determined based on the preferences of the candidate  $w$  and Condition 9 of Definition 1).
4. Verify that the number of students ranking the candidate  $c'_j$  is exactly as required in the conditions ( $2\bar{q}$  for  $j > 1$ , and  $q_{c'_1}^1 + q_{c'_1}^0 - 1$  for the candidate  $c'_1$ ) and that all of them rank the funded contract with  $c'_j$  first and the unfunded contract with  $c'_j$  second. If this isn't the case, declare that the set is empty and stop. Otherwise, set these students as candidate  $s_i^j$ , with indices determined by Conditions 2 and 3 of Definition 1.
5. Declare that the event indexed by the vector of candidates was realized.

It is clear that if an event of the form  $E^n(c, z)$  is realized the algorithm must identify it. Since the output of the algorithm is a unique event, this proves that  $E^n(c, z)$  and  $E^n(c, z')$  are disjoint for different selections of  $z$  and  $z'$ , as required.  $\square$

**Lemma 5.** For any student  $w$ , and any two different selections,  $z$  and  $z'$ , of

$$\left( c, \{c'_j\}_{j=1}^k, p, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \{\{s_i^j\}_{i=1}^{2\bar{q}}\}_{j=2}^k \right),$$

the events  $E^n(w, z)$  and  $E^n(w, z')$  are disjoint.<sup>29</sup>

*Proof.* Towards contradiction, assume that both events are realized. By Condition 9 of Definition 1,  $z$  and  $z'$  share the same  $c'_1$  (it must be a party to the highest ranked funded contract according to  $w$ 's preferences). Hence,  $z$  and  $z'$  share the same  $p$  (it is the only student that ranks the funded contract with  $c'_1$  second after another funded contract. This implies that  $z$  and  $z'$  share the same  $c$  (it is the college that is a party to the first ranked contract of  $p$ ). But this is a contradiction to Lemma 4, that states that events that share the same selection of  $c$  are disjoint.  $\square$

**Lemma 6.** For any student  $p$ , and any two different selections,  $z$  and  $z'$ , of

$$\left( c, \{c'_j\}_{j=1}^k, w, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \{\{s_i^j\}_{i=1}^{2\bar{q}}\}_{j=2}^k \right),$$

the events  $E^n(p, z)$  and  $E^n(p, z')$  are disjoint.

<sup>29</sup>We abuse notation for readability. The notation is intended to represent a student playing the role of  $w$  in Definition 1. The same remark applies to Lemma 6.

*Proof.* By Condition 8 of Definition 1, if both events are realized, then  $z$  and  $z'$  share the same  $c$  (it must be a party to the highest ranked contract according to  $p$ 's preferences). But this is a contradiction to Lemma 4.  $\square$

**Lemma 7.** For any student  $r_1$ , and any two different selections,  $z$  and  $z'$ , of

$$\left( c, \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=2}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \left\{ \{s_i^j\}_{i=1}^{2\bar{q}} \right\}_{j=2}^k \right),$$

the events  $E^n(p, z)$  and  $E^n(p, z')$  are disjoint.

*Proof.* By Condition 7 of Definition 1, if both events are realized, then  $z$  and  $z'$  share the same  $c$  (it must be a party to the highest ranked contract according to  $r_1$ 's preferences). But this is a contradiction to Lemma 4.  $\square$

*Proof of Theorem 3.* First, note that by Lemmas 1 and 3 the expected number of events  $E^n(\cdot)$  that are realized is at least  $L \cdot n$  for some  $L \in (0, 1)$ . Second, Lemma 4 guarantees that fixing  $c$ , the events  $E^n(c, z)$  are disjoint for different selections of  $z$  (where  $z$  is a shorthand for all other selections needed to an event as in Definition 1). This means that the probability that fewer than  $Ln/2$  colleges play the role of  $c$  in an event  $E^n(c, z)$  that is realized is at most  $(2 - 2L)/(2 - L)$ . In other words, the probability that more than  $Ln/2$  colleges play the role of  $c$  in an event  $E^n(c, z)$  that is realized is at least  $L/(2 - L)$ .

Third, note that if the preference-flip algorithm is initialized to include in  $A'$  only the contracts of the form  $(r_1, c, 1)$  for events  $E^n \left( c, \{c'_j\}_{j=1}^k, p, w, \{r_i\}_{i=1}^{q_c^1}, \{s_i^1\}_{i=1}^{q_{c'_1}^1 + q_{c'_1}^0 - 1}, \left\{ \{s_i^j\}_{i=1}^{2\bar{q}} \right\}_{j=2}^k \right)$  that are realized, the resulting allocation differs from  $\Upsilon^{SP-DA}$  only in that students playing the role of  $r_1$  will get  $(r_1, c, 0)$  instead of  $(r_1, c, 1)$ , students playing the role of  $p$  will be assigned  $(p, c, 1)$  instead of  $(p, c'_1, 1)$ , and students playing the role of  $w$  will be assigned  $(w, c'_1, 0)$  instead of being unassigned. Lemma 5 guarantees that  $w$ 's are different across these events. Hence, this algorithm never decreases the number of assigned students, and with probability at least  $L/(2 - L)$ , it increases the number of assigned students by a factor of, at least,  $L/2\lambda$ . Lemmas 6 and 7 provide a similar guarantee for the upward transfers and funding losers.

Finally, we note that it is possible to initialize the algorithm in polynomial time (e.g., by applying the algorithm identified in Lemma 4  $n$  times, once for each college), and that our algorithm will terminate in polynomial time (indeed, after the initialization, it will run SP-DA once, and it is well known that the running-time of SP-DA is polynomial).  $\square$

**Remark.** Our proof relies only on the probability with which the events from Definition 1 are realized. So long as this probability remains sufficiently large (at least for a large fraction of these events) our proof will go through. Examples include modes of randomization that restrict *substitutes* to rank funded positions higher than unfunded positions as well as modes of randomization with correlation structures as in Kojima and Pathak (2009), where some schools are more popular than others, or students' preferences are drawn from heterogeneous distributions (e.g., some students prefer economics majors while others prefer physics).



The proofs of Proposition 5, Proposition 6, and Proposition 7 are provided in Supplemental Appendix B.

## **Supplemental Material (for Online Publication)**

## Appendix B Supplemental Proofs (for Online Publication)

*Proof of Theorem 2.* We will prove that Theorem 2 holds even if colleges are restricted to offer no more than one seat under each of the financial terms. Our proof relies on a reduction—we show that an algorithm that identifies a stable (certainly stable) allocation of maximum size quickly (in running time polynomial in the size of the input) can also be used to quickly produce a solution to another problem that is known to be NP-hard. Therefore, identifying such an algorithm will prove that  $P = NP$ .

The NP-hard problem that we reduce is a special version of MAX-SMTI, studied in [Manlove et al. \(2002\)](#). The description of the problem is as follows. An instance of *the restricted stable marriage problem with incomplete lists and ties* consists of a set of  $n$  men,  $M$ , and a set of  $n$  women  $W = W^\sim \cup W^\infty$ , where men in  $M$  and women in  $W^\infty$  have strict preferences over agents on the other side and remaining unmatched, and women in  $W^\sim$  are indifferent between two acceptable men and find all other men unacceptable. Given such an instance,  $I$ , a matching is *weakly stable* if no agent's assignment is unacceptable and there is no pair of agents that strictly prefer one another to their assignment. The problem of deciding whether  $I$  admits a stable matching under which all women are matched with men is NP-complete ([Manlove et al., 2002](#)).

Given an instance  $I = \langle M, W^\sim, W^\infty, \{\succeq_i\}_{i \in M \cup W} \rangle$  we define *the corresponding Hungarian college admissions market* as follows. The set of students is  $S^I = \{s_i\}_{i \in M \cup W^\sim}$ , with elements corresponding to each man and each woman in  $W^\sim$ . The set of colleges  $C^I = \{c_w\}_{w \in W^\sim \cup W^\infty}$ . For each  $w \in W^\infty$ ,  $c_w$  has a single state-funded seat ( $q_{c_w}^1 = 1$  and  $q_{c_w}^0 = 0$ ), while other colleges have one state-funded seat and one self-funded seat (i.e., for all  $w \in W^\sim$ ,  $q_{c_w}^1 = 1$  and  $q_{c_w}^0 = 1$ ).

For each  $w \in W^\infty$ ,  $c_w$  ranks students (and the outside option) according to the preferences of the corresponding woman ( $s_m \gg_{c_w} s_{m'}$  iff  $w$  prefers  $m$  to  $m'$ ,  $s_m \gg_{c_w} \emptyset$  iff  $w$  prefers  $m$  to her outside option, and  $\emptyset \gg_{c_w} s_{w'}$  for all  $w' \in W^\sim$ ). For each  $w \in W^\sim$ ,  $c_w$  has three acceptable students  $s_w \gg_{c_w} s_{p_w} \gg_{c_w} s_{f_w}$  where  $p_w$  and  $f_w$  are the two men acceptable to  $w$  (ranked according to some arbitrary rule). Furthermore, for each  $w \in W^\sim$ ,  $s_w$ 's most preferred alternative is the state-funded seat in  $c_w$  followed by a self-funded seat in this college, with all other alternatives being unacceptable ( $(s_w, c_w, 1) \succ_{s_w} (s_w, c_w, 0) \succ_{s_w} \emptyset$ , and  $\emptyset \succ_{s_w} (s_w, c, t)$  for all  $c \neq c_w$  and all  $t \in \{0, 1\}$ ).

Finally, for each  $m \in M$ ,  $s_m$  ranks contracts according to  $m$ 's preferences (ranking the state-funded alternative over self-funded one) except for contracts corresponding to  $w \in W^\sim$  where only one of the terms is acceptable (depending on the students ranking on  $c_w$ 's ranking). Formally,  $s_m$ 's preferences satisfy the following conditions:

1. for each  $w \in W^\sim$ ,  $(s_m, c_w, 1) \succ_{s_m} \emptyset$  only if  $s_m$  is ranked second according to  $\gg_{c_w}$ , and  $(s_m, c_w, 0) \succ_{s_m} \emptyset$  only if  $s_m$  is ranked third according to  $\gg_{c_w}$ .
2. for each  $w \in W$  and  $t$  such that the contracts  $(s_m, c_w, t)$  is not unacceptable by the first condition  $(s_m, c_w, t) \succ_{s_m} \emptyset$  iff  $w$  is acceptable to  $m$ .
3. for each  $w, w' \in W$  and  $t, t'$  such that the contracts  $(s_m, c_w, t)$  and  $(s_m, c_{w'}, t')$  are not unacceptable by the first condition,  $(s_m, c_w, t) \succ_{s_m} (s_m, c_{w'}, t')$  iff  $m$  prefers  $w$  to  $w'$ .

**Lemma 8.** *An instance  $I$  of the restricted stable marriage problem with incomplete lists and ties admits a stable matching under which all women are matched with men if and only if the corresponding Hungarian college admissions market admits a stable allocation of cardinality  $|S^I|$  (i.e., where all students are assigned to colleges).*

*Proof.* Assume  $I$  admits a stable matching  $\mu$  under which all women are matched with men. Consider the allocation  $Y$  consisting of the contracts  $\left\{ \left( s_{\mu(w)}, c_w, 1 \right) \right\}_{w \in W^\sim}$  together with the contracts  $\left\{ \left( s_{\mu(w)}, c_w, t_{\mu(w)} \right) \right\}_{w \in W^\sim}$  and  $\left\{ \left( s_w, c_w, 1 - t_{\mu(w)} \right) \right\}_{w \in W^\sim}$ , where  $t_{\mu(w)}$  guarantees that  $\left( s_{\mu(w)}, c_w, t_{\mu(w)} \right) \succ_{s_{\mu(w)}} \emptyset$  for each  $w \in W^\sim$ .<sup>30</sup> Then  $|Y| = |S^I|$  and  $Y$  is a (certainly) stable allocation.

The only non-trivial case to verify is that of potential blocks involving colleges with two seats (corresponding to women in  $W^\sim$ ). This case follows since for each of these colleges the student involved in  $\left( s_{\mu(w)}, c_w, t_{\mu(w)} \right) \succ_{\mu(w)} \emptyset$  finds matching with the college under the other financial terms unacceptable. Hence, if the college is assigned its first- and second-ranked students, the college does not prefer any other allocation that is individually rational. Furthermore, if the college is assigned its first- and third-ranked students, then the first ranked student receives state funding (her most preferred contract), and the only individually rational allocation that the college prefers would require her to receive a less preferred alternative (the self-funded seat).

In the other direction, let  $Y$  be a (certainly) stable allocation such that  $|Y| = |S^I|$ . For each  $w \in W^\sim$  the student  $s_w$  must be assigned to  $c_w$  (under some financial terms). This holds since  $w$  ranks  $s_w$  first, and  $s_w$  only ranks the contracts with  $c_w$  as acceptable. Furthermore, since  $S^I$  is equal to the number of available seats, all seats are assigned.

The matching  $\nu$  that assigns  $w$  to  $m$  if and only if  $(\{s_m\} \times \{c_w\} \times T) \cap Y \neq \emptyset$  is individually rational, and it clearly matches every women to a men.<sup>31</sup> We claim that it is also weakly stable. To see this, we note that no  $w \in W^\sim$  can be involved in a blocking pair, as these women get their (tied) first choice men. But a blocking involving  $w \in W^\sim$  and  $m \in M$  implies that  $Y$  is blocked through  $\{(s_m, c_w, 1)\}$ .  $\square$

**Proposition 8.** *Finding a maximum-size stable (or certainly stable) allocation in Hungarian college admissions markets is NP-hard, even when colleges offer no more than one seat under each of the financial terms (i.e.,  $q_c^t \leq 1$  for each  $c$  and  $t$ ).*

*Proof.* Transforming the restricted stable marriage problem with incomplete lists and ties to the corresponding Hungarian college admissions market clearly requires only a polynomial running time. An algorithm that finds a maximum-size stable (or certainly stable) allocation in Hungarian college admissions markets should find a stable allocation that matches all students if and only if one exists. If the algorithm is guaranteed to output a maximum-size stable allocation in polynomial time, transforming the original marriage problem and

<sup>30</sup>Existence of such financial terms is guaranteed since  $w$  must be acceptable to  $\mu(w)$  and vice versa, by the stability of  $\mu$ .

<sup>31</sup>That  $\nu$  is indeed a matching follows by the feasibility of  $Y$  and the fact that in any stable allocation colleges with more than one seat must be a side to a contracts with a student  $s_w$  for some  $w \in W^\sim$ .

running the algorithm will give an answer in polynomial time to whether the original marriage problem admits a stable matching under which all women are matched with men (by Lemma 8). And this problem is NP-complete (Manlove et al., 2002).  $\square$

This completes the proof of Theorem 2.  $\square$

**Remark.** MAX-SMTI is not approximable within a factor of 21/19, unless  $P = NP$  (Halldórsson et al., 2007). Using this fact and the construction used to prove Lemma 8, one can establish that the maximum-size stable (or certainly stable) allocation in Hungarian college admissions markets is not approximable within a constant factor 15/14, even when colleges offer no more than one seat under each of the financial terms (i.e.,  $q_c^t \leq 1$  for every  $c$  and  $t$ ).

## Proof of Proposition 5

**Lemma 9.** *Let  $Y$  be a stable allocation in a Hungarian college admissions market where all students are not sensitive to funding. Consider a modified market that is identical except that for each college,  $c$ , quotas are given by  $\hat{q}_c^0 = q_c^0 + q_c^1$  and  $\hat{q}_c^1 = 0$ . Then, the allocation  $\{(s, c, 0) \mid \exists t \in T, \text{ s.t. } (s, c, t) \in Y\}$  is stable in the modified market.*

*Proof.* The modified market is (effectively) a matching without contracts market with responsive preferences. Hence, it is sufficient to consider blocking by pairs and violations of individual rationality (Proposition 1). Individual rationality is guaranteed by the stability of  $Y$  and students not being sensitive to funding. Similarly, if the matching (allocation) is blocked by the pair  $s$  and  $c$  (through  $\{(s, c, 0)\}$ ), then  $Y$  is blocked through  $\{(s, c, 0)\}$  or  $\{(s, c, 1)\}$ .  $\square$

*Proof of Proposition 5.* By Lemma 9, each stable allocation  $Y$  has the same number of students assigned to each college as the number of students that are assigned to the corresponding college in a stable matching of a certain matching market without contracts where colleges preferences are responsive. The statement therefore follows from the rural hospital theorem.  $\square$

## Proof of Proposition 6

**Lemma 10.** *If students rank at most one contract with each college, an allocation is stable if and only if it corresponds to a stable matching in the related market.*

*Proof.* The “if” direction follows from Theorem 1. For the “only if” direction, note that if the matching is not individually rational, so is the corresponding allocation. Furthermore, if an individually rational matching is blocked, it must be blocked by a pair  $(c, t)$  and  $s$ , such that  $s$  is not matched to  $(c, 1 - t)$  (by individual rationality) nor to  $(c, t)$  (by the definition of blocking). Therefore, the corresponding allocation is blocked through the contract  $(s, c, t)$ .  $\square$

**Lemma 11.** *Let  $Y$  be a stable allocation in  $\langle S, C, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C}, \{\succ_s\}_{s \in S} \rangle$ , and let  $B$  denote the set of students that consider two contracts with the same college acceptable. Then each  $s' \in S \setminus B$*

weakly prefers  $Y^{SP-DA} \left[ \left\langle S \setminus B, C, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C}, \{\succ_s\}_{s \in S \setminus B} \right\rangle \right]$  to  $Y$ , and prefers  $Y$  to the allocation  $Y^{SR-DA} \left[ \left\langle S \setminus B, C, \{\succ_c, \gg_c, \max\{0, q_c^t - |\{s \in B \mid (s, c, t) \succ_s \emptyset\}|\}\}_{c \in C, t \in \{0,1\}}, \{\succ_s\}_{s \in S \setminus B} \right\rangle \right]$ . Furthermore, the opposite comparisons hold for colleges that are not ranked by students in  $B$ .

*Proof.* Consider  $\left\langle S \setminus B, C, \{\succ_c, \gg_c, \max\{0, q_c^t - |\{s \in B \mid (s, c, t) \succ_s \emptyset\}|\}\}_{c \in C, t \in \{0,1\}}, \{\succ_s\}_{s \in S \setminus B} \right\rangle$ , the market where students in  $B$  are removed and the quotas are reduced at contracts they rank.  $Y \upharpoonright_{S \setminus B}$  must be stable in this market (otherwise,  $Y$  is not stable). Furthermore, since students in  $S \setminus B$  do not rank more than one contract with the same college, the allocation must also be stable in the related matching market (by Lemma 10). Hence, by the consensus property, all students in  $S \setminus B$  prefer the outcome of SP-DA in this market to  $Y \upharpoonright_{S \setminus B}$  to the outcome of SR-DA (and the opposite holds for colleges not ranked by students in  $B$ ). Furthermore, by Theorem 2.25 in Roth and Sotomayor (1990), all students in  $S \setminus B$  prefer the outcome of SR-DA in  $\left\langle S \setminus B, C, \{\succ_c, \gg_c, q_c^0, q_c^1\}_{c \in C}, \{\succ_s\}_{s \in S \setminus B} \right\rangle$  to the outcome of SR-DA in  $\left\langle S \setminus B, C, \{\succ_c, \gg_c, \max\{0, q_c^t - |\{s \in B \mid (s, c, t) \succ_s \emptyset\}|\}\}_{c \in C, t \in \{0,1\}}, \{\succ_s\}_{s \in S \setminus B} \right\rangle$  (and the opposite holds for colleges not ranked by students in  $B$ ).  $\square$

**Lemma 12 (Kojima and Pathak).** Consider a matching (without contracts) market and a college  $c$  that is matched with  $\mu(c)$  under SP-DA. If for any non-empty subset of  $\mu(c)$ , dropping this subset does not result in an additional offer to  $c$  relative to SP-DA, then  $c$  is matched with  $\mu(c)$  also under SR-DA.

**Randomized  $(\hat{c}, \hat{t})$  Algorithm** Building on Lemmas 11 and 12, we present a randomized algorithm that declares failure with probability higher than the probability that there are stable allocations  $Y$  and  $Y'$  such that  $Y \cap S \times \{\hat{c}\} \times \{\hat{t}\} \neq Y' \cap S \times \{\hat{c}\} \times \{\hat{t}\}$  (i.e., different sets of agents are assigned to  $\hat{c}$  under  $\hat{t}$  in different stable allocations). Roughly, the algorithm randomly draws the identities and preferences of students in  $B$  (the set defined in Lemma 11) and fails if  $|B|$  is unusually large or if one of its members ranks a contract with  $\hat{c}$ . If the algorithm doesn't fail, agents in  $B$  are removed and the algorithm proceeds to mimic the distribution of outcomes of SP-DA, where students preferences are drawn as needed based on the distribution that conditions on them not being members of  $B$ . Next, the algorithm transitions from the outcome of SP-DA in this market, to a market where quotas are reduced as in Lemma 11. The algorithm declares failure if  $\hat{c}$  receives an offer in the process (a necessary condition for the assignment of  $(\hat{c}, \hat{t})$  to change) or if the process induces an "unusually long" chain of offers.

**Step 1. Drawing the set  $B$  and their preferences:** For each student in  $S$  draw preferences according to the *independent* mode of randomization. Denote by  $B$  the (random) set of students who rank two contracts with the same college. If a student in  $B$  ranks a contract with  $\hat{c}$ , or if  $|B| > \left(1 + \log^2(|C|)\right) \cdot |S| \cdot 8k^2/|C|$  declare failure and stop. Otherwise, "forget" the preferences of students in  $S \setminus B$  and continue to Step 2.

**Step 2. SP-DA excluding  $B$ :**



Step 2.1 For each  $s \in S \setminus B$ , draw independently and uniformly at random a college from  $C$ , and draw independently and uniformly at random terms from  $T = \{0, 1\}$  with equal probability for 0 and 1. Have each of these students propose the auxiliary college (college-terms pair) she drew. Auxiliary colleges accept temporarily the most preferred subset of students and reject all others.

Step 2.j>1 If all students in  $S \setminus B$  are temporarily assigned, or have been rejected  $2k$  times, set  $Y^1$  to be the resulting allocation and proceed to Step 3. Otherwise, for each student that is not temporarily assigned and has not been rejected  $2k$  times, draw independently and uniformly at random a college in  $C$  to which the students did not apply in a previous step of the algorithm (under any terms), and random terms from  $T = \{0, 1\}$  with equal probability for 0 and 1. Have each student propose the auxiliary college she drew. Auxiliary colleges accept temporarily the most preferred subset of the union of the set of new proposal and the proposals they temporarily hold and reject all others. Proceed to Step 2.j + 1.

Step 3. **from SP-DA excluding  $B$  to SP-DA excluding  $B$  with reduced quotas** Set  $Y = Y^1$ .

Let  $(s_m, c_m, t_m)_{m=1}^{2k|B|}$  be the set of contracts ranked by students in  $B$ . For each  $m$  from 1 to  $2k|B|$ :

Step 3.m Reduce  $q_{c_m}^{t_m}$  by 1 (unless it is already 0). If this does not result in  $c_m$  violating its quota under  $Y$ , proceed to step 3.m + 1. Otherwise, reject the lowest ranked student,  $s$ , assigned to the auxiliary college  $(c_m, t_m)$ , and proceed to Step 3.m.1.

Step 3.m.j Denote by  $P_1 > 0$  a parameter (we will specify a value for  $P_1$  at a later point in the proof (see Footnote 32)). If  $j > P_1 \cdot \log n$  declare failure and stop. If  $s$  has been rejected  $2k$  times, proceed to step 3.m + 1. Otherwise draw independently and uniformly at random a college to which the student did not apply in a previous step of the algorithm,  $c$ , and random terms from  $t \in \{0, 1\}$  with equal probability for 0 and 1. If  $c = \hat{c}$  declare failure and stop. Have  $s$  propose to the corresponding auxiliary college,  $(c, t)$ . The auxiliary college accepts temporarily the most preferred subset of  $(Y \cap (S \times \{c\} \times \{t\})) \cup \{(s, c, t)\}$ , and  $Y$  is updated accordingly to  $\text{Ch}_{(c,t)}(Y \cup \{(s, c, t)\}) \cup (Y \setminus (S \times \{c\} \times \{t\}))$ . If the auxiliary college rejects some student (at most one students can be rejected), set  $s$  to be this student, and proceed to Step 3.m.j + 1. Otherwise, proceed to Step 3.m + 1.

Step 3.(2k|B|+1) Set  $Y^2$  to be the resulting allocation. proceed to Step 4.

Step 4. **Stochastic rejection chains:** We note that  $|Y \cap (S \times \{\hat{c}\} \times \{\hat{t}\})| \leq \bar{q}$ , and thus this set has no more than  $2^{\bar{q}}$  different subsets. For each nonempty subset  $Z = \{(s_i, \hat{c}, \hat{t})\}_{i=1}^{l_Z}$  of  $Y \cap (S \times \{\hat{c}\} \times \{\hat{t}\})$ :

Step 4.Z Set  $Y = Y^2$  and realized preferences to the part that was drawn by the end of Step 3. Have  $(\hat{c}, \hat{t})$  drop all students in  $Z_S$  from its preference list (at most  $\bar{q}$  students) and reject them from the college. Set  $i = b = 1$ , and proceed as follows:

Step 4.Z.i.b If  $i = l_Z + 1$ , proceed to the next subset. If  $b > P_1 \log n$  declare failure and stop. If  $s_i$  has been rejected  $2k$  times, move to step 4.Z.i + 1.1. Otherwise draw indepen-

dently and uniformly at random a college to which the student did not apply in a previous step of the algorithm,  $c$ , and random terms  $t \in \{0, 1\}$  with equal probability for 0 and 1. If the college drawn is  $\hat{c}$ , declare failure and stop. Otherwise let the chosen auxiliary college,  $(c, t)$ , choose from  $Y \cup \{(s_i, c, t)\}$ .

If  $(c, t)$  does not reject any student, move to Step 4.Z.i + 1. Otherwise, set  $s_i$  to be the rejected student. Update  $Y$  according to the choice of  $(c, t)$ . Proceed to Step 4.Z.i.b + 1.

**Termination** After all possible subsets of  $Z$  (fewer than  $2^{\bar{q}}$ ) have been exhausted, stop.

**Lemma 13.** *For any  $(\hat{c}, \hat{t}) \in C^n \times T$ , the probability that  $(\hat{c}, \hat{t})$  has a unique assignment in all stable allocations is bounded below by the probability that the  $(\hat{c}, \hat{t})$  randomized algorithm terminates without declaring failure.*

*Proof.* By construction:

- The probability that the algorithm fails in Step 1 is equal to the probability that the set  $B$  has certain properties (regardless of whether or not  $(\hat{c}, \hat{t})$  has a unique assignment in all stable allocations).
- The probability that the algorithm fails in Step 3 conditional on not failing in Step 1 is equal to the probability – conditional on  $B$  not having the aforementioned properties – that reducing the quotas changes the SP-DA assignment of  $(\hat{c}, \hat{t})$  or leads to long rejection chains.
- The probability that the algorithm fails in Step 4 conditional on not failing in earlier step is greater than the probability – conditional on  $B$  not having the aforementioned properties and quota reductions not changing the SP-DA assignment of  $(\hat{c}, \hat{t})$  or leading to long rejection chains – that the SP-DA and SR-DA assignments of  $(\hat{c}, \hat{t})$  under reduced quotas are equal (using Boole’s inequality and Lemma 12).

The result therefore follows by Lemma 11. □

**Lemma 14** (Typical Size of  $B$ ). *Let  $\{\tilde{\Gamma}^n\}_{n=1}^\infty$  be a regular sequence of markets such that the mode of randomization is independent, and let  $\tilde{B}^n$  denote the set of students who rank two contract with the same college. Then the probability that  $|\tilde{B}^n| > (1 + \log^2(n)) \times |S^n| \times 8k^2/n$  is lower than  $1/n^2$  for sufficiently large  $n$ .*

*Proof.* When  $n > 4k$  the probability that an agent finds more than one contract with some college acceptable is

$$1 - 1 \times \left(1 - \frac{1}{2n-1}\right) \times \left(1 - \frac{2}{2n-2}\right) \times \cdots \times \left(1 - \frac{2k-1}{2n-(2k-1)}\right) \leq 1 - \left(1 - \frac{4k}{n}\right)^{2k}$$

which, by Bernoulli’s inequality, is bounded above by

$$1 - \left(1 - \frac{8k^2}{n}\right) = \frac{8k^2}{n}.$$

Since students' preferences are drawn independently of one another, the events in which one of them finds more than one contract with the same college acceptable are independent. Hence, using a multiplicative Chernoff bound, this implies that the probability that more than  $(1 + \log^2(n)) \times |S^n| \times 8k^2/n$  students find more than one contract with the same college acceptable is lower than

$$\exp\left(-\log^2(n) \cdot 4k^2 \cdot |S^n|/n\right) \leq \exp\left(-\log^2(n) \cdot 4\lambda k^2\right),$$

and

$$\lim_{n \rightarrow \infty} n^2 \cdot \exp\left(-\log^2(n) \cdot 4\lambda k^2\right) = 0.$$

□

**Lemma 15.** *Given a regular sequence of markets, there exists  $n'$  and  $M > 0$  such that for all  $n > n'$  and any college  $\hat{c} \in C^n$ , the probability that a contract with  $\hat{c}$  is ranked by any student in  $B^n$  is bounded above by  $M \log^2(n)/n$ .*

*Proof.* Unless  $B$  is larger than  $(1 + \log^2(n)) \cdot |S^n| \cdot 8k^2/n$  (which, for large  $n$ , occurs with probability lower than  $1/n^2$  by Lemma 14) the number of contracts ranked by agents in  $B$  is lower than  $2k \cdot (1 + \log^2(n)) \cdot |S^n| \cdot 8k^2/n$ . Each of these contracts has probability of  $1/n$  to be with  $\hat{c}$ . The result follows by Boole's inequality and the fact that  $\log^2(n)/n > 1/n^2$  for large  $n$ . □

### The Effect of Quotas Reduction on SP-DA

**Lemma 16 (Immorlica and Mahdian).** *Given a regular sequence of markets, there exists  $\alpha > 0$  such that for any  $n$ , the probability that more than  $\alpha n$  colleges in  $C^n$  are not acceptable to any student in  $S \setminus B$  is greater than  $1 - 1/n^2$  under both the independent and the substitutes modes of randomization.*

*Proof.* The probability that a particular college is ranked by some student is smaller than  $1 - (1 - k/n)^{n/\lambda}$ . Therefore, the expected number of colleges ranked by some student is lower than  $M(n) := n \left(1 - (1 - k/n)^{n/\lambda}\right) \geq n(1 - \exp(-k/\lambda))$ . The indicators of these events meet the conditions of Theorem 1.1 of Impagliazzo and Kabanets (2010) (they are negatively correlated), and therefore the probability that more than  $(1 + \epsilon)M(n)$  of them are realized is bounded above by  $\exp(-2n\epsilon^2)$ . □

*Proof of Proposition 6.* By Lemma 13, the probability that the Randomized  $(\hat{c}, \hat{t})$  Algorithm fails is greater than the probability with which  $Y \cap (S \times \{\hat{c}\} \times \{\hat{t}\})$  varies for different choices of a stable allocation  $Y$ . By Lemmas 14 and 15 the probability that it fails in Step 1 is lower than  $A \cdot \log^2(n)/n$  for sufficiently large  $A > 0$ . Conditional on not failing in Step 1 the arguments of Immorlica and Mahdian (and Kojima and Pathak) show that the probability

that the algorithm fails is bounded above by  $A' \cdot \log(n)/n$  for sufficiently large  $A' > 0$ .<sup>32</sup> Furthermore, the probability that a student has multiple stable assignments is lower than the probability that she ranks a contract with a college that has multiple stable assignments, which is lower than  $2k$  times the probability that a college has multiple stable assignments.

Finally, let  $X_{(c,t)}$  denote the random variable that the set of students assigned to college  $c$  under  $t$  differs across stable allocations. By the discussion above, there exists  $\Delta > 0$  such that  $\mathbb{E} \left[ \sum_{(c,t) \in C \times T} X_{(c,t)} \right] < \Delta \log^2 n$ . Hence, by Markov's inequality

$$\Pr \left\{ \sum_{(c,t) \in C \times T} X_{(c,t)} > \Delta \sqrt{n} \log^2 n \right\} < \frac{1}{\sqrt{n}}.$$

The proof for students is completely analogous. □

**Remark.** Our proof continues to hold under conditions that guarantee that [Immorlica and Mahdian's](#) short rejection chains argument remains valid. Examples include the variant of the *independent* mode of randomization where students systematically prefer state-funded seats to self-funded seats in the same program as well as modes where students' preferences are drawn from heterogeneous distributions (e.g., some students prefer economics majors while others prefer physics).

## Proof of Proposition 7

*Sketch of Proof of Proposition 7.* First, relying on Proposition 2, we consider the related markets. Second, we note that under the *independent* mode of randomization, the related markets meet the conditions of [Kojima and Pathak \(2009\)](#).

Next, we consider the *substitutes* mode of randomization. This case is slightly more complicated. The added complexity arises from the fact that it is likely that when an auxiliary college rejects a student, this student will displace another student at the "sibling" auxiliary college, who will go on to propose to the auxiliary college—resulting in an additional proposal. We note, however, that Lemma 12 can be strengthened, to require that the additional offer to the (auxiliary) college comes from a student that is preferred to all rejected students.

---

<sup>32</sup>This argument is now standard in the literature (e.g. [Immorlica and Mahdian, 2005](#); [Kojima and Pathak, 2009](#); [Kojima et al., 2013](#); [Ashlagi et al., 2014](#)) and so we omit it for brevity. To summarize: by Lemma 16 there exists excess capacity in more than  $\alpha n$  colleges with probability greater than  $1 - 1/n^2$ . The probability of failure due to long rejection chain is bounded above by the sum of the probability that there is no such excess capacity and the probability that a geometric random variable reaches values larger than  $P_1 \log(n)$  for a constant  $P_1 > 0$  (conditional on not stopping, the probability of a chain to stop in the next step continues to be at least  $\alpha$ ). The probability of a geometric random variable to exceed  $\log(n)/\alpha$  steps is

$$(1 - \alpha)^{\log(n)/\alpha} = \left( (1 - \alpha)^{1/\alpha} \right)^{\log(n)} < \frac{1}{e} = \frac{1}{n}$$

In the algorithm, we choose the parameter  $P_1 \geq 1/\alpha$ , allowing us to invoke this bound. Conditional on no failure due to a long rejection chain, the probability of failure (since an offer to  $c$  was made) is lower than  $2P_1 \log(n)/n$ .

Furthermore, students who get displaced from the “sibling” auxiliary college by strategically rejected students are not ranked sufficiently high to meet this condition themselves (as otherwise, they would not be displaced) and indirect chains of rejection are unlikely as in the argument of [Immorlica and Mahdian \(2005\)](#) and [Kojima and Pathak \(2009\)](#) (see Lemma 16 and Footnote 32 for details of the argument).  $\square$

**Remark.** The proof extends to modes of randomization with correlation structures as in [Kojima and Pathak \(2009\)](#), where some schools are more popular than others, or students’ preferences are drawn from heterogeneous distributions (e.g., some students prefer economics majors while others prefer physics). The proof for modes of randomization that restrict *independent* or *substitutes* to rank funded positions higher than unfunded positions is simpler, as it rules out the complication of short rejection cycles discussed above.

## Appendix C Supplemental Tables (for Online Publication)

Section C.1 compares the realized and the benchmark assignments. Section C.2 presents additional summary statistics.

### C.1 Data

This Appendix presents the number of applicants assigned in 2007 and in our benchmark assignment (Table C1).

Table C1: Realized and benchmark assignments

	Realized (1)	Benchmark (2)
Assigned to a contract	81,563	84,130
Assigned to a state-funded contract	48,726	48,725
Assigned to a self-funded contract	32,837	35,405

*Notes:* The table presents the number of applicants assigned under each of the financial terms in the realized assignment in 2007 (column (1)) and in our benchmark assignment (column (2)).

### C.2 Summary Statistics

This appendix presents additional summary statistics. Table C2 summarizes the means and standard deviations of the applicants' background characteristics. Table C3 provides summary statistics on the characteristics of applicants' ROLs by the type of the settlement where they reside – an additional proxy for socioeconomic status (cf. Table 2).

Tables C4 and C5 present the coefficients of a linear regression of ROL characteristics on disadvantaged status and 11th-grade GPA on the sample of applicants. We find that disadvantaged applicants are more (less) likely to rank state-funded (self-funded) contracts exclusively in their ROLs, and they are less likely to rank both contracts with the same study program consecutively. We also find that these differences cannot be explained by the differences in applicants' academic achievement, and thus, by the differences in their admission chances.

Next, we consider three alternative proxies for socioeconomic status: per-capita annual gross income, NABC-based SES index, and the type of the settlement where applicants reside (capital, county capital, town, and village). Tables C6 and C7 show that conditional on academic achievement, applicants of higher socioeconomic status are more likely to rank at least one self-funded contract in their ROL and they are more likely to rank two contracts with the same study program consecutively.



Table C2: Summary statistics on applicants' characteristics

	Mean (1)	SD (2)	N (3)
Disadvantaged	0.05	0.21	108,854
Per-capita annual gross income (1000 USD, 2007 prices)	9.90	2.30	106,934
NABC-based SES index	0.22	0.37	84,455
Capital	0.20	0.40	106,934
County capital	0.21	0.41	106,934
Town	0.33	0.47	106,934
Village	0.26	0.44	106,934
11th-grade GPA (1-5)	3.63	0.84	85,811
Female	0.57	0.50	108,854
Number of alternatives in ROL	3.71	2.21	108,854
Number of programs in ROL (observed)	3.01	1.30	108,854

*Notes:* The table reports mean values and standard deviations of applicant characteristics. Disadvantaged status is an indicator for claiming priority points for disadvantaged status. 11th-grade GPA is the average grades in mathematics, history, and Hungarian grammar and literature. Applicants' settlement of residence is missing for 1,920 applicants, NABC-based SES index is missing for 24,399 applicants, and 11th-grade GPA is missing for 23,043 applicants.

Table C3: Summary statistics on applicants' ROLs: Applicants' residence

	All applicants (%) (1)	Capital (%) (2)	County town (%) (3)	Town (%) (4)	Village (%) (5)
<i>A. Preference for funding</i>					
State-funded contract exclusively	51.4	37.1	52.5	54.3	58.2
Self-funded contract exclusively	18.5	24.0	18.8	17.5	15.2
State- and self-funded contracts	30.1	38.9	28.7	28.2	26.6
Same study program consecutively	15.5	20.5	14.6	14.3	13.5
Same study program consecutively on the top of the ROL	11.6	15.0	11.3	10.7	10.2
<i>B. Preference for study characteristics</i>					
Single program location	49.4	64.5	50.0	44.4	43.5
Single university	35.0	28.8	43.2	33.7	34.7
Single faculty	26.3	24.4	30.4	25.2	25.8
Single field of study	54.9	58.9	55.2	54.1	52.4
Single major	29.6	33.5	30.5	28.7	26.9
# of applicants	108,854	21,731	22,381	35,290	27,532

*Notes:* The table reports summary statistics on applicants' ROL. Panel A focuses on preference for funding. Specifically, Panel A shows the share of applicants who rank state-funded (self-funded) contracts exclusively, the share of students who rank state-funded and self-funded contracts as well, and the share of applicants who rank the same study program with state-funding and self-funding consecutively, and who rank the same study program with state-funding and self-funding consecutively on the top of the ROL. Panel B focuses on preference for program characteristics. Specifically, Panel B shows the share of applicants who rank exclusively contracts that are in the same settlement (single program location), at a single university (single university), and at a single faculty of a university (single university), and the share of applicants who rank exclusively contracts in a single field of study, and in a single major. Column (1) presents these shares for all applicants, and columns (2)–(5) report these shares for applicants residing in the capital, county towns, towns, and villages, respectively.

Table C4: Socioeconomic status, academic achievement, and ROL characteristics

Dependent variable	Ranked state-funded contract exclusively			Ranked self-funded contract exclusively		
	(1)	(2)	(3)	(4)	(5)	(6)
Disadvantaged	0.294*** (0.006)		0.277*** (0.006)	-0.172*** (0.002)		-0.152*** (0.002)
11th-grade GPA (standardized)		0.079*** (0.002)	0.079*** (0.002)		-0.032*** (0.001)	-0.032*** (0.001)
Mean outcome (non-disadvd.)	0.500	0.500	0.500	0.192	0.192	0.192
R-squared	0.015	0.035	0.049	0.009	0.047	0.054

*Notes:* The table presents the coefficient of linear regressions of ROL characteristics (such as whether an applicant ranked state-funded contracts exclusively on her ROL (columns (1)–(3)) and whether an applicant ranked self-funded contracts exclusively on her ROL (columns (4)–(6))) on socioeconomic status and academic achievement. The sample includes 108,854 applicants. The regressions include indicators for missing values of standardized 11th-grade GPA (23,043 applicants). Robust standard errors are in parentheses.

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

Table C5: Socioeconomic status, academic achievement, and ROL characteristics (2)

Dependent variable	Same study program consecutively			Same study program consecutively on the top of the ROL		
	(1)	(2)	(3)	(4)	(5)	(6)
Disadvantaged	-0.075*** (0.004)		-0.075*** (0.004)	-0.057*** (0.004)		-0.056*** (0.004)
11th-grade GPA (standardized)		-0.034*** (0.001)	-0.034*** (0.001)		-0.027*** (0.001)	-0.027*** (0.001)
Mean outcome (non-disadvd.)	0.158	0.158	0.158	0.119	0.119	0.119
R-squared	0.002	0.007	0.009	0.001	0.006	0.007

*Notes:* The table presents the coefficient of linear regressions of ROL characteristics (such as whether an applicant ranked the same study program consecutively (columns (1)–(3)) and whether an applicant ranked the same study program consecutively on the top of the ROL (columns (4)–(6))) on socioeconomic status and academic achievement. The sample includes 108,854 applicants. The regressions include indicators for missing values of standardized 11th-grade GPA (23,043 applicants). Robust standard errors are in parentheses.

\*\*\*:  $p < 0.01$ , \*\*:  $p < 0.05$ , \*:  $p < 0.1$ .

Table C6: Socioeconomic status, academic achievement, and ROL characteristics: Additional specifications

Dependent variable	Ranked state-funded contract exclusively			Ranked self-funded contract exclusively		
	(1)	(2)	(3)	(4)	(5)	(6)
Per-capita annual gross income (1000 USD, 2007 prices)	-0.028*** (0.001)			0.013*** (0.001)		
NABC-based SES index		-0.098*** (0.005)			-0.010*** (0.003)	
Capital			-0.208*** (0.002)			0.078*** (0.002)
County capital			-0.061*** (0.003)			0.035*** (0.003)
Town			-0.043*** (0.003)			0.023*** (0.003)
11th-grade GPA (standardized)	0.081*** (0.002)	0.077*** (0.002)	0.081*** (0.002)	-0.033*** (0.001)	-0.017*** (0.001)	-0.033*** (0.001)
R-squared	0.057	0.088	0.056	0.055	0.145	0.053

*Notes:* The table presents the coefficient of linear regressions of ROL characteristics (such as whether an applicant ranked state-funded contracts exclusively on her ROL (columns (1)–(3)) and whether an applicant ranked self-funded contracts exclusively on her ROL (columns (4)–(6))) on socioeconomic status and academic achievement. The sample includes 108,854 applicants. The regressions include indicators for missing values of standardized 11th-grade GPA (23,043 applicants), per-capita annual gross income (1,920 applicants), NABC-based SES index (24,399 applicants), and residence (1,920 applicants). The omitted category in columns (3) and (6) is village. Robust standard errors are in parentheses.

\*\*\*:  $p < 0.01$ , \*\*:  $p < 0.05$ , \*:  $p < 0.1$ .

Table C7: Socioeconomic status, academic achievement, and ROL characteristics: Additional specifications (2)

Dependent variable	Same study program consecutively			Same study program consecutively on the top of the ROL		
	(1)	(2)	(3)	(4)	(5)	(6)
Per-capita annual gross income (1000 USD, 2007 prices)	0.009*** (0.000)			0.009*** (0.000)		
NABC-based SES index		0.061*** (0.004)			0.044*** (0.003)	
Capital			0.073*** (0.004)			0.050*** (0.004)
County capital			0.014*** (0.003)			0.014*** (0.003)
Town			0.010*** (0.002)			0.007*** (0.002)
11th-grade GPA (standardized)	-0.035*** (0.001)	-0.041*** (0.001)	-0.035*** (0.001)	-0.028*** (0.001)	-0.031*** (0.001)	-0.028*** (0.001)
R-squared	0.011	0.010	0.013	0.010	0.008	0.009

*Notes:* The table presents the coefficient of linear regressions of ROL characteristics (such as whether an applicant ranked the same study program consecutively (columns (1)–(3)) and whether an applicant ranked the same study program consecutively on the top of the ROL (columns (4)–(6))) on socioeconomic status and academic achievement. The sample includes 108,854 applicants. The regressions include indicators for missing values of standardized 11th-grade GPA (23,043 applicants), per-capita annual gross income (1,920 applicants), NABC-based SES index (24,399 applicants), and residence (1,920 applicants). The omitted category in columns (3) and (6) is village. Robust standard errors are in parentheses.

\*\*\*:  $p < 0.01$ , \*\*:  $p < 0.05$ , \*:  $p < 0.1$ .



Table C8: Program characteristics by excess self-funded capacity at the benchmark

	Programs with no excess self-funded capacity		Programs with excess self-funded capacity		Programs with excess self-funded capacity (weighted)	
	(1)	(2)	(3)	(4)	(5)	(6)
Periphery	0.59	(0.49)	0.85	(0.36)	0.84	(0.37)
Full-time	0.90	(0.30)	0.72	(0.45)	0.71	(0.45)
Field: economics and business	0.09	(0.28)	0.10	(0.30)	0.20	(0.40)
Field: humanities	0.15	(0.36)	0.16	(0.37)	0.11	(0.32)
Field: engineering	0.05	(0.21)	0.09	(0.29)	0.09	(0.28)
Field: natural science	0.05	(0.21)	0.05	(0.21)	0.02	(0.15)
Field: social science	0.08	(0.27)	0.09	(0.29)	0.09	(0.29)
State-funded priority-score cutoff (scale: 0–144)	104.65	(21.17)	100.37	(20.96)	102.39	(20.97)
Excess self-funded seats per program	0.00	(0.00)	29.53	(33.27)	29.53	(33.27)
# number of programs	494	-	941	-	941	-

*Notes:* The table presents means and standard deviations of various characteristics of programs with and without excess self-funded capacity at the benchmark. Columns (1)–(4) use equal weights, while columns (5)–(6) weight programs by the number of excess self-funded seats at the benchmark. Standard deviations are in parentheses.

## Appendix D Exercising Market Power More Broadly and More Narrowly

This appendix expands our empirical analysis in two directions. First, we assess the consequences of being able to detect only a subset of students over which market power is applied by the preference flip algorithm. Second, we present alternate algorithms that apply market power more broadly and analyze their consequences.

### D.1 Exercising Market Power More Narrowly

We consider the robustness of our results on the preference flip algorithm to cases where market power is applied more narrowly. Specifically, we apply market power only over a subset of applicants and rule out the possibility to apply market power over other applicants even if they rank two contracts with the same program consecutively. This analysis is motivated by the possibility that information frictions prevent the designer from identifying the entire set  $A$  (from the description of the preference flip algorithm).

Appendix Table D1 presents the number of assigned applicants when market power is applied more narrowly, and the increase in the number of assigned applicants relative to the benchmark SR-DA assignment. We find that the increase in the number of assigned applicants is roughly linear in the fraction of applicants over which market power can be applied.

Table D1: Narrow market power

Fraction subject to market power (%)	Assigned under preference flip	Increase relative to SR-DA (%)
(1)	(2)	(3)
100	85,688	1.9
90	85,538	1.7
80	85,391	1.5
70	85,236	1.3
60	85,079	1.1
50	84,929	0.9
40	84,774	0.8
30	84,609	0.6
20	84,454	0.4
10	84,293	0.2

*Notes:* The table presents the number of assigned applicants under the preference flip algorithm, when market power is applied only on a subset of applicants. In each row, we consider a fraction of applicants over which market power can be applied (from 10 to 100 percent), and rule out the possibility to apply market power over other applicants even if they rank two contracts with the same program consecutively. We present the average result over 50 random draws of subsets of this size. Column 3 presents the increase in the number of assigned students relative to the benchmark (84,130).

### D.2 Exercising Market Power More Broadly

In this section, we present two additional stable algorithms that can be interpreted as applying market power more broadly. We then repeat the analysis from the main text using these

algorithms instead of the Preferences Flip Algorithm.

**Greedy Reject Algorithm** Initialize a set  $B' \subseteq B$ , where  $B$  is the set of student–college pairs such that the student is assigned to a state-funded seat in the program under SP-DA. For each pair in  $B'$ , remove from the student’s initial ROL the state-funded contract with the college (the student’s assignment) and run SP-DA on the resulting problem. If the resulting allocation is certainly stable (with respect to original preferences), stop and output the resulting allocation. Otherwise, remove some pairs from  $B'$  and repeat the process.

**Combined Algorithm** Initialize sets  $A' \subseteq A$  and  $B' \subseteq B$ . For each pair in  $A'$ , flip the order of the contracts with the college in the student’s original ROL, so that the state-funded contract appears immediately *after* the self-funded contract. For each pair in  $B'$ , remove from the student’s initial ROL the state-funded contract with the college (the student’s assignment). Run SP-DA on the resulting problem. If the resulting allocation is certainly stable (with respect to original preferences), stop and output the resulting allocation. Otherwise, remove some pairs from  $A'$  or from  $B'$  and repeat the process.

**Proposition 9.** *For any initial set(s)  $B' \subseteq B$  (and  $A' \subseteq A$ ) and any rule regarding the choice of elements to remove from  $B'$  (and  $A'$ ), the greedy reject (combined) algorithm results in a stable allocation. Furthermore, the resulting allocation may not allocate funding based on merit.*

*Proof.* The proof is identical to the proof of Proposition 4 and therefore omitted. □

We now repeat the analysis from the main text using the greedy reject algorithm and the combined algorithm instead of the preference flip algorithm. Appendix Table D2 compares the outcomes of the greedy reject and the combined algorithms to the SR-DA benchmark. Appendix Tables D3 and D4 compare the characteristics of winners to losers under the greedy reject algorithm and the combined algorithm, respectively. Appendix Tables D5 and D6 investigate geographic mobility under the greedy reject algorithm and the combined algorithm, respectively. The results are qualitatively similar to those from the main text, but their magnitudes are slightly larger.

Table D2: Winners, losers, and the number of assigned applicants: Greedy reject and combined algorithms

	Benchmark SR-DA (1)	SP-DA (2)	Greedy reject algorithm (3)	Combined algorithm (4)
<i>A. Number of assigned applicants</i>				
Assigned to a contract	84,130	84,130	85,575	85,847
Assigned to a state-funded contract	48,725	48,725	48,710	48,706
Assigned to a self-funded contract	35,405	35,405	36,865	37,141
Assigned to a self-funded, full-time contract	13,321	13,321	14,410	14,565
<i>B. Winners</i>				
		8	5,056	5,283
– Newly assigned		0	1,969	2,346
– New program		8	2,168	2,498
– Same program, preferred financial terms		0	919	439
<i>C. Losers</i>				
		0	4,570	4,734
– Newly unassigned		0	524	629
– New program		0	645	762
– Same program, less preferred financial terms		0	3,401	3,343

*Notes:* The table presents the number of assigned applicants under various stable assignments (Panel A). Panels B and C describe the number of winners and losers from changing the benchmark SR-DA to SP-DA and to our alternate algorithms. An applicant is a winner if she is assigned to a contract she ranked higher than her contract in the benchmark assignment. An applicant is a loser if she is assigned to a contract she ranked lower than her assigned contract in the benchmark assignment or if she becomes unassigned. Column (1) shows the benchmark assignment. Column (2) presents the SP-DA assignment. Column (3) presents the greedy reject algorithm assignment and column (4) presents the combined algorithm assignment.

Table D3: Characteristics of winners and losers: Greedy reject algorithm

	Winners (1)	Losers (2)	p-values: (1)=(2) (3)
Disadvantaged	0.057	0.026	0.000
Per-capita annual gross income (1000 USD, 2007 prices)	9.735	10.385	0.000
NABC-based SES index	0.180	0.306	0.000
Capital	0.188	0.277	0.000
County capital	0.192	0.217	0.003
Town	0.329	0.298	0.001
Village	0.291	0.208	0.000
11th-grade GPA (1-5)	3.473	3.704	0.000
Female	0.562	0.553	0.355
Number of applicants	5,056	4,570	9,626

*Notes:* The table reports the mean values of characteristics of winners and losers from changing the benchmark SR-DA to the greedy reject algorithm. An applicant is a winner if she is assigned to a contract she ranked higher than her benchmark assignment. An applicant is a loser if she is assigned to a contract she ranked lower than her assigned contract in the benchmark assignment or if she becomes unassigned. Column (3) presents p-values for the equality of mean characteristics of winners and losers.

Table D4: Characteristics of winners and losers: Combined algorithm

	Winners (1)	Losers (2)	p-values: (1)=(2) (3)
Disadvantaged	0.061	0.026	0.000
Per-capita annual gross income (1000 USD, 2007 prices)	9.718	10.413	0.000
NABC-based SES index	0.177	0.305	0.000
Capital	0.183	0.282	0.000
County capital	0.191	0.214	0.005
Town	0.331	0.297	0.000
Village	0.295	0.207	0.000
11th-grade GPA (1-5)	3.483	3.689	0.000
Female	0.570	0.556	0.160
Number of applicants	5,283	4,734	10,017

*Notes:* The table reports the mean values of characteristics of winners and losers from changing the benchmark SR-DA to the combined algorithm. An applicant is a winner if she is assigned to a contract she ranked higher than her benchmark assignment. An applicant is a loser if she is assigned to a contract she ranked lower than her assigned contract in the benchmark assignment or if she becomes unassigned. Column (3) presents p-values for the equality of mean characteristics of winners and losers.

Table D5: Geographic mobility: Greedy reject algorithm

	Mover		Assigned to periphery		Moved to capital	
	SR-DA (1)	GRA (2)	SR-DA (3)	GRA (4)	SR-DA (5)	GRA (6)
<i>A. Total</i>						
All assigned applicants	51,221	52,249	47,845	48,876	22,444	22,804
<i>B. Winners</i>						
Newly assigned (N= 1,969)	–	1,249	–	1,362	–	402
New program (N= 2,168)	1,379	1,415	1,323	1,147	521	660
<i>C. Losers</i>						
Newly unassigned (N= 524)	279	–	226	–	143	–
New program (N= 645)	348	370	256	327	191	153

*Notes:* The table presents the number of movers (i.e., applicants whose assigned study program is not located in the county where they reside), of applicants assigned to the periphery (i.e., not to the capital), and of applicants who moved to the capital (i.e., applicants who get assigned to a study program in the capital but do not reside there) under the greedy reject algorithm (GRA) and in the benchmark SR-DA assignments. Panel A presents the total number of applicants with these characteristics. Panels B and C focus on winners and losers who are assigned to a different study program.



Table D6: Geographic mobility: Combined algorithm

	Mover		Assigned to periphery		Moved to capital	
	SR-DA (1)	CA (2)	SR-DA (3)	CA (4)	SR-DA (5)	CA (6)
<i>A. Total</i>						
All assigned applicants	51,221	52,415	47,845	49,073	22,444	22,851
<i>B. Winners</i>						
Newly assigned (N= 2,346)	–	1,463	–	1,620	–	472
New program (N= 2,498)	1,582	1,624	1,526	1,316	602	769
<i>C. Losers</i>						
Newly unassigned (N= 629)	341	–	265	–	183	–
New program (N= 762)	419	449	294	377	232	183

*Notes:* The table presents the number of movers (i.e., applicants whose assigned study program is not located in the county where they reside), of applicants assigned to the periphery (i.e., not to the capital), and of applicants who moved to the capital (i.e., applicants who get assigned to a study program in the capital but do not reside there) under the combined algorithm and in the benchmark SR-DA assignments. Panel A presents the total number of applicants with these characteristics. Panels B and C focus on winners and losers who are assigned to a different study program.

## Appendix E Robustness Analysis (for Online Publication)

This Appendix shows that our empirical findings are robust to the way we handle minor inconsistencies in our data. Our data contain 3,686 applicants (3.4 percent of all applicants) whose reported assignment and reported priority scores are inconsistent with stability (their reported priority score does not exceed the priority-score cutoff of their reported assignment, or there is an alternative in their ROL from which they are rejected even though their reported priority score exceeds the priority-score cutoff of this alternative). These inconsistencies largely stem from applicants having a reported priority score of zero. Our main approach, which is presented in the text, holds the assignment of applicants with such inconsistencies fixed to their reported assignment, and makes the corresponding seats unavailable to others.

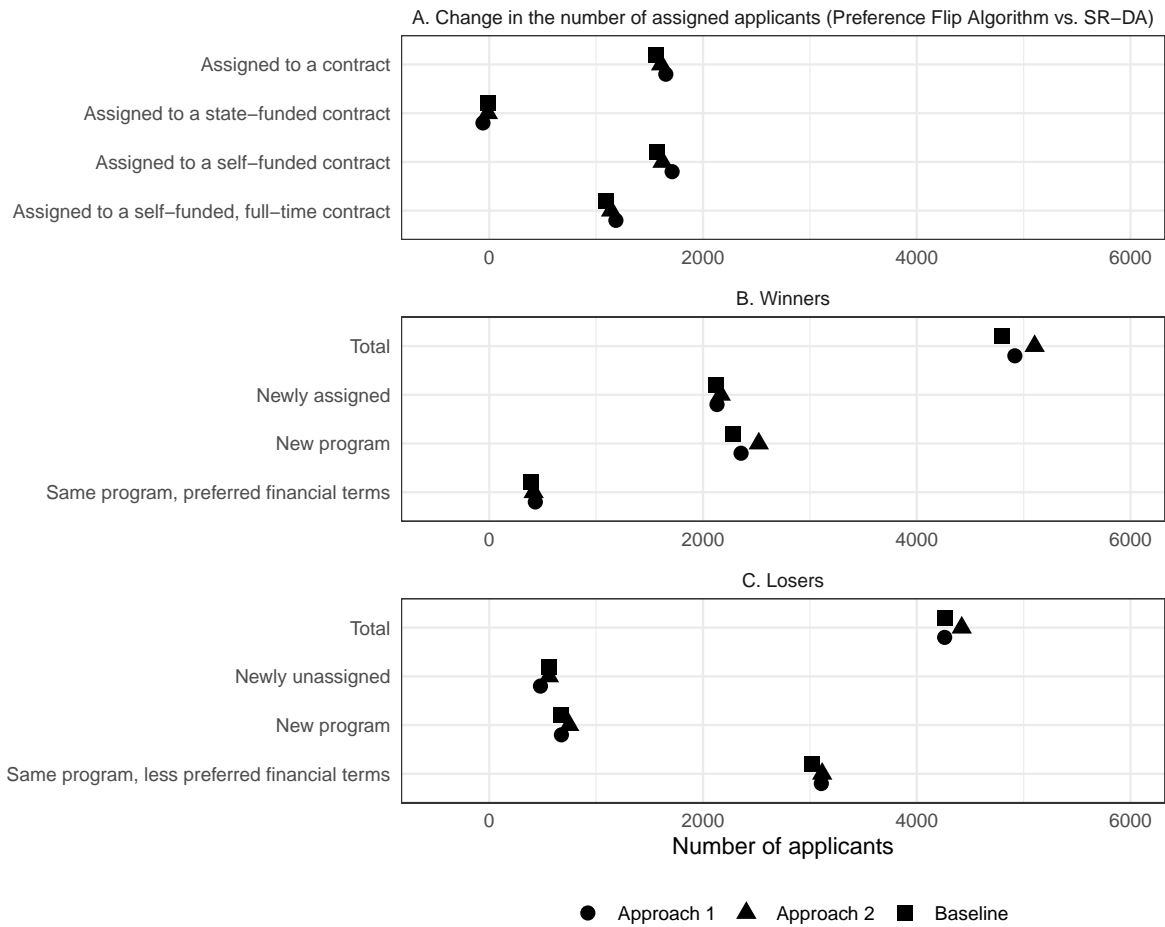
In this appendix, we consider two alternative approaches to verify that this issue does not drive our empirical findings:

- Approach 1: We hold the assignment of applicants with inconsistencies fixed to their reported assignment, but we keep the corresponding seats available to others (i.e., we do not reduce the corresponding quotas)
- Approach 2: We do not hold the assignment of applicants with inconsistencies fixed, and we keep the corresponding seats available to others.

Approach 1 changes the SR-DA (benchmark) assignment of 4,439 applicants and increases the number of assigned applicants in the benchmark by 1,989 relative to our main approach. Approach 2 changes the SR-DA (benchmark) assignment of 6,052 applicants and increases the number of assigned applicants in the benchmark by 162 relative to our main approach. In spite of these changes, Figure E1 shows that our main findings continue to hold.

Panel A of Figure E1 presents the difference between the number of assigned applicants under the preference flip algorithm and under SR-DA. In the main text, this corresponds to the difference between column (3) and column (1) in Panel A of Table 3. Panel B presents the number of winners from changing the benchmark to the preference flip algorithm. In the main text, this corresponds to column (3) in Panel B of Table 3. Panel C presents the number of losers from changing the benchmark to the preference flip algorithm. In the main text, this corresponds to column (3) in Panel C of Table 3. Since the numbers in each of the panels are almost identical, we conclude that our main findings continue to hold.

Figure E1: Winners, losers, and the number of assigned applicants: Robustness



*Notes:* This figure studies the robustness of our empirical findings to the way we handle minor inconsistencies in our data. Panel A presents the difference between the number of assigned applicants under the preference flip algorithm and under SR-DA. In the main text (baseline), this corresponds to the difference between column (3) and column (1) in Panel A of Table 3. Panel B presents the number of winners from changing the benchmark to the preference flip algorithm. In the main text, this corresponds to column (3) in Panel B of Table 3. Panel C presents the number of losers from changing the benchmark to the preference flip algorithm. In the main text, this corresponds to column (3) in Panel C of Table 3. The baseline corresponds to the analysis in the main text.

## Appendix F Variable Description (for Online Publication)

This appendix describes the construction of variables in Table C2.

- **Disadvantaged status:** Applicants receive priority points for having a low socioeconomic background. Our administrative data report these priority points. If an applicant received priority points for this reason in any of the alternatives in her ROL, we label this applicant as disadvantaged. Source: Administrative data on college admissions.
- **NABC-based SES index:** First, we compute the NABC-based SES index for each survey respondent between 2008 and 2012 (5 years) in grade 10. We then compute the high-school-specific means over the 2008–2012 period. Finally, we merge the high-school-specific NABC-based SES index with the applicants using their high-school identifier. The high-school identifier is missing for 15,521 applicants. Source: National Assessment of Basic Competencies (NABC).
- **Per-capita annual gross income:** Settlement-level income per population. Applicants come from 2,804 different settlements. Average exchange rate in 2007: 183HUF/USD. Source: T-STAR dataset ([http://adatbank.krtk.mta.hu/adatbazisok\\_\\_\\_tstar](http://adatbank.krtk.mta.hu/adatbazisok___tstar))
- **Capital, county capital, town, village:** Dummy variables for the type of settlement where the applicants reside. Source: Administrative data on college admissions.
- **11th-grade GPA:** Average grade in mathematics, Hungarian grammar and literature, and history in grade 11. Source: Administrative data on college admissions.
- **Female:** Dummy variable for being female. Source: Administrative data on college admissions.
- **Number of alternatives in ROL:** Our administrative data include information on the number of alternatives in each applicant's ROL. Source: Administrative data on college admissions.
- **Number of programs in ROL (observed):** Our administrative data report the first 6 alternatives in an applicant's ROL as well as the applicant's realized assignment, in case it was ranked lower. We compute the number of programs in an applicant's ROL based on this information. Source: Administrative data on college admissions.

## Appendix G Evidence from the Israeli Psychology Master’s Match (for Online Publication)

In this Appendix, we evaluate our theoretical predictions using data from the Israeli Psychology Master’s Match (IPMM). In this market, unlike in Hungary, applicants are not heterogeneous in the sensitivity to financial terms. Consistent with our theory, we find that the preference flip algorithm does not increase the number of applicants admitted to college relative to the outcome of SP-DA.

### G.1 Background and Data

We use data from the 2014 and 2015 rounds of the IPMM. In each of these years, there are approximately 1,000 applicants, of which approximately 600 are assigned.

Several departments offer a limited number of no-strings-attached scholarships to selected students. In 2014, 10 programs in 3 departments offered admission under multiple financial terms. The number increased to 15 programs in 4 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. The centralized clearinghouse uses a version of SP-DA which allows programs to offer contracts with multiple financial terms, and allows applicants to rank these alternatives separately.<sup>33</sup>

### G.2 The Distribution of Applicants’ Rank-Order Lists

Applicants’ ROLs include 4.32 contracts on average ( $SD=4.14$ ). About 37.2% of the ROLs include at least one contract with a program that offered admission under multiple financial terms. Of these, only 3.4% include only the funded contract in some program but not the unfunded contract (cf. Table 2).

Among the ROLs that include both a funded and an unfunded contract with the same program, more than 90% ranked the funded contract first. Among these ROLs, 82.6% ranked both contracts consecutively, and the mean number of contracts ranked between a funded contract and the corresponding unfunded contract was 0.34.

In summary, an overwhelming majority of applicants in the IPMM are not sensitive to financial terms.

### G.3 The Preference Flip Allocation

We next compare the stable allocations resulting from SP-DA and from the preference flip algorithm.<sup>34</sup> We implement the preference flip algorithm initializing  $A' = A$  and with multiple rules of removal: in addition to the rule used in the main analysis, we also remove the lowest-ranked student (instead of the highest-ranked student), and furthermore we use 100

---

<sup>33</sup>While colleges’ preferences in the IPMM are more complex than in Hungarian college admissions (Hassidim et al., 2017b), they meet the hidden substitutes condition of Hatfield and Kominers (2015) and so SP-DA is stable and strategyproof.

<sup>34</sup>The proof that the preference flip algorithm results in a stable allocation continues to hold in the more general preference structure of the IPMM (since SP-DA continues to be stable).

randomly generated rules of removal. Consistent with Proposition 5, none of these implementations result in an increase in the number of assigned students, and the largest number of affected students is 4.