Incentives and Stability in Large Two-Sided Matching Markets *

Fuhito Kojima† Parag A. Pathak‡

January 2009

Abstract

A number of labor markets and student placement systems can be modeled as many-to-one matching markets. We analyze the scope for manipulation in many-to-one matching markets under the student-optimal stable mechanism when the number of participants is large. Under some regularity conditions, we show that the fraction of participants that have incentives to misrepresent their preferences when others are truthful approaches zero as the market becomes large. With an additional technical condition, truthful reporting by every participant is an approximate equilibrium under the student-optimal stable mechanism in large markets. The results help explain the success of the student-optimal stable mechanism in large matching markets observed in practice.

Keywords: Large markets, stability, two-sided matching.
JEL Class: C78, D61, D78.

*We appreciate helpful discussions with Attila Ambrus, Eric Budish, Federico Echenique, Drew Fudenberg, Akihiko Matsui, Alvin E. Roth, Tayfun Sönmez, Satoru Takahashi and seminar participants at Harvard, Hitotsubashi, Kyoto, Osaka, Tokyo, 2006 North American Summer Meetings of the Econometric Society at Minneapolis, the 17th International Conference on Game Theory at SUNY Stony Brook and 2006 INFORMS Annual Meeting at Pittsburgh. Discussions with Michihiro Kandori and Yohei Sekiguchi were especially important for the current version of the paper. Detailed comments by the co-editor and anonymous referees greatly improved the substance and exposition of the paper. For financial support, Pathak is grateful to the National Science Foundation, the Spencer Foundation, and the Division of Research at Harvard Business School.

†Cowles Foundation, Yale University, New Haven, CT 06510, e-mail: fuhitokojima1979@gmail.com.

‡Society of Fellows, Harvard University, Cambridge, MA 02138 and Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02142, e-mail: ppathak@fas.harvard.edu.

In recent years, game theoretic ideas have been used to study the design of markets. Auctions have been employed to allocate radio spectrum, timber, electricity, and natural gas involving hundreds of billions of dollars worldwide (Milgrom (2004)). Matching procedures have found practical applications in centralized labor markets, as well as school assignment systems in New York City and Boston.[1] Connections between auctions and matching have been explored and extended (Kelso and Crawford (1982), Hatfield and Milgrom (2005), Day and Milgrom (2007)).

The practical aspects of market design have led to an enrichment of the theory of market design. This paper investigates a theoretical problem motivated by the use of stable matching mechanisms in large markets, inspired by a practical issue first investigated by Roth and Peranson (1999).[2] A matching is stable if there is no individual agent who prefers to become unmatched or pair of agents who prefer to be assigned to each other to being assigned their allocation in the matching. In real-world applications, empirical studies have shown that stable mechanisms often succeed whereas unstable ones often fail.[3]

Although stable mechanisms have a number of virtues, they are susceptible to various types of strategic behavior before and during the match. Roth (1982) shows that any stable mechanism is manipulable via preference lists: for some participants, reporting a true preference list (ordinal ranking over potential matches) may not be a best response to reported preferences of others. In many-to-one markets such as matching markets between colleges and students, where some colleges have more than one position, Sönmez (1997a) and Sönmez (1999) show that there are additional strategic concerns. First, any stable mechanism is manipulable via capacities, that is, colleges may sometimes benefit by underreporting their quotas. Second, any stable mechanism is manipulable via pre-arranged matches; that is, a college and a student may benefit by agreeing to match before receiving their allocations from the centralized matching mechanism.

Concerns about the potential for these types of manipulation are often present in real markets. For instance, in New York City (NYC) where the Department of Education has recently adopted a stable mechanism, the Deputy Chancellor of Schools described principals concealing capacity as a major issue with their previous unstable mechanism (New York Times (11/19/04)):

> "Before you might have a situation where a school was going to take 100 new children for 9th grade, they might have declared only 40 seats, and then placed the

[1]For a survey of this theory, see Roth and Sotomayor (1990). For applications to labor markets, see Roth (1984a) and Roth and Peranson (1999). For applications to student assignment, see for example Abdulkadiroğlu and Sönmez (2003), Abdulkadiroğlu, Pathak, Roth and Sönmez (2005) and Abdulkadiroğlu, Pathak and Roth (2005).

[2]A recent paper by Bulow and Levin (2006) theoretically investigates a matching market with price competition, motivated by an antitrust case against the National Residency Matching Program. See also Niederle (2007).

[3]For a summary of this evidence, see Roth (2002).

other 60 outside of the process."[4]

Roth and Rothblum (1999) discuss similar anecdotes about preference manipulation from the National Resident Matching Program (NRMP), which is an entry-level matching market for hospitals and medical school graduates in the U.S.

The aim of this paper is to understand why despite these negative results many stable mechanisms appear to work well in practice. We focus on the student-optimal stable mechanism, which forms the basis of many matching mechanisms used in the field such as in NYC and the NRMP. Dubins and Freedman (1981) and Roth (1982) show that students do not have incentives to manipulate markets with the student-optimal stable mechanism, though colleges do have incentives to do so. Our results show that the mechanism is immune to various kinds of manipulations by colleges when the market is large. In real-world two-sided matching markets, there are often a large number of applicants and institutions, and each applicant submits a preference list containing only a small fraction of institutions in the market. For instance, in the NRMP, the length of the applicant preference list is about 15, while the number of hospital programs is between 3,000 and 4,000 and the number of students is over 20,000 per year. In NYC, about 75% of students submit preference lists of less than 12 schools, and there are about 500 school programs and over 90,000 students per year.[5] Motivated by these features, we study how the scope for manipulations changes when the number of market participants grows but the length of the preference lists does not.

We consider many-to-one matching markets with the student-optimal stable mechanism, where colleges have arbitrary preferences such that every student is acceptable, and students have random preferences of fixed length drawn iteratively from an arbitrary distribution. We show that the expected proportion of colleges that have incentives to manipulate the student-optimal stable mechanism when every other college is truth-telling converges to zero as the number of colleges approaches infinity. The key step of the proof involves showing that, when there are a large number of colleges, the chain reaction caused by a college's strategic rejection of a student is unlikely to make another more preferred student apply to that college.[6]

[4]A careful reader may recognize that this quote is not about a stable mechanism but about an unstable one. We present this quote just to suggest that strategic behavior may be a realistic problem in general. Indeed, concern about strategic behavior motivated New York City's recent adoption of a stable mechanism, and this paper suggests that scope for manipulations may be limited under the current stable mechanism despite theoretical possibility of manipulation.

[5]For data regarding the NRMP, see http://www.nrmp.org/2006advdata.pdf. For data regarding New York City high school match, see Abdulkadiroğlu, Pathak and Roth (2008).

[6]Throughout the paper we focus on the SOSM. The college-optimal stable mechanism can be similarly defined by letting colleges propose to students. By making complementary assumptions about the model, we can derive results concerning incentives of students under the college-optimal stable mechanism similar to those in the current paper. However, additional assumptions may be needed for analyzing the incentives of colleges in this case, since truth-telling is not a dominant strategy for colleges with a quota larger than one in the college-optimal stable mechanism.

The above result does not necessarily mean that agents report true preferences *in equilibrium*. Thus we also conduct equilibrium analysis in the large market. We introduce an additional technical condition, called sufficient thickness of the markets, and show that truthful reporting is an approximate equilibrium in a large market that is sufficiently thick.

Related literature

Our paper is most closely related to Roth and Peranson (1999) and Immorlica and Mahdian (2005). Roth and Peranson (1999) conduct a series of simulations on data from the NRMP and on randomly generated data. In their simulations, very few agents could have benefited by submitting false preference lists or by manipulating capacity in large markets when every other agent is truthful. These simulations lead them to conjecture that the fraction of participants with preference lists of limited length who can manipulate tends to zero as the size of the market grows.[7]

Immorlica and Mahdian (2005), which this paper builds upon, theoretically investigate one-to-one matching markets where each college has only one position and show that as the size of the market becomes large, the proportion of colleges that are matched to the same student in all stable matchings approaches one. Since a college does not have incentives to manipulate via preference lists if and only if it is matched to the same student in all stable matchings (see Gale and Sotomayor (1985) and Demange, Gale and Sotomayor (1987)), this result implies that most colleges cannot manipulate via preference lists when the market is large.

While this paper is motivated by these previous studies, there are a number of crucial differences. First, our focus in this paper is on many-to-one markets, which include several real-world markets such as the NRMP and the school choice program in NYC. In such markets a college can sometimes manipulate via preference lists even if the college is matched to the same set of students in all stable matchings.[8] Moreover, in many-to-one markets there exists the additional possibilities of capacity manipulation and manipulation via pre-arrangement (unraveling) which are not present in a one-to-one market.[9] We introduce new techniques to overcome these complications.

Second, previous research mostly focuses on counting the average number of participants that can manipulate the student-optimal stable mechanism, assuming that others report their preferences truthfully. This leaves open the question of whether participants will behave truth-

[7]The property that in large matching problems the size of the set of stable matchings has also been documented using data from Boston Public School's assignment system. Pathak and Sönmez (2007) report that in Boston, for the first two years of data from SOSM, in elementary school, the student-optimal stable matching and college-optimal stable matching coincide, while for middle school, at most three students are assigned to different schools in the two matchings.

[8]For instance, see the example in Theorem 5.10 of Roth and Sotomayor (1990).

[9]Indeed, Roth and Peranson (1999) explicitly investigate the potential for capacity manipulation in their simulations.

fully *in equilibrium*. A substantial part of this paper investigates this question and shows that truthful reporting is an approximate equilibrium in a large market that is sufficiently thick.[10]

Incentive properties in large markets are studied in other areas of economics. For example, in the context of double auctions, Gresik and Satterthwaite (1989), Rustichini, Satterthwaite and Williams (1994), Pesendorfer and Swinkels (2000), Swinkels (2001), Fudenberg, Mobius and Szeidl (2007), and Cripps and Swinkels (2006) show that the equilibrium behavior converges to truth-telling as the number of traders increases under various informational structures. Papers more closely related to ours are discussed in Section 4.

There is a literature that analyzes the consequences of manipulations via preference lists and capacities in complete information finite matching markets. See Roth (1984b), Roth (1985) and Sönmez (1997b) for games involving preference manipulation and Konishi and Ünver (2006) and Kojima (2006) for games of capacity manipulation. Some of the papers most relevant to ours are discussed in Section 4.

The paper proceeds as follows. Section 1 presents the model and introduces a lemma which is key to our analysis. Section 2 defines a large market and presents our main result. Section 3 conducts equilibrium analysis. Section 4 concludes. All proofs are in the Appendix.

1 Model

1.1 Preliminary definitions

Let there be a set of students S and a set of colleges C. Each student s has a strict preference relation P_s over the set of colleges and being unmatched (being unmatched is denoted by s). Each college c has a strict preference relation $\succ_c$ over the set of subsets of students. If $s \succ_c \emptyset$, then s is said to be **acceptable** to c. Similarly, c is acceptable to s if $cP_s s$. A market is tuple $\Gamma = (S, C, P_S, \succ_C)$ where $P_S = (P_s)_{s \in S}, \succ_C = (\succ_c)_{c \in C}$.

Since only rankings of acceptable mates matter for our analysis, we often write only acceptable mates to denote preferences. For example,

$$P_s : c_1, c_2,$$

means that student s prefers college c_1 most, then c_2, and c_1 and c_2 are the only acceptable colleges.

For each college $c \in C$ and any positive integer q_c, its preference relation $\succ_c$ is **responsive with quota** q_c if the ranking of a student is independent of her colleagues, and any set of

[10] Immorlica and Mahdian (2005) claim that truth-telling is an approximate equilibrium in a one-to-one market even without sufficient thickness. In section 4, we present an example to show that this is not the case, but truth-telling is an approximate equilibrium under an additional assumption of sufficient thickness.

students exceeding quota q_c is unacceptable (see Roth (1985) for a formal definition).[11] We will assume that all preferences are responsive throughout the paper.

Let P_c be the corresponding **preference list of college** c, which is the preference relation over individual students and $\emptyset$. Sometimes only the preference list structure and quotas are relevant for the analysis. We therefore sometimes abuse notation and denote by $\Gamma = (S, C, P, q)$ an arbitrary market in which the preferences induce preference lists $P = (P_i)_{i \in S \cup C}$ and quotas $q = (q_c)_{c \in C}$. We say that $(S, C, P_S, \succ_C)$ induces (S, C, P, q) in such a case. We also use the following notation: $P_{-i} = (P_j)_{j \in S \cup C \setminus i}$ and $q_{-c} = (q_{c'})_{c' \in C \setminus c}$.

A **matching** μ is a mapping from $C \cup S$ to itself such that (i) for every s, $|\mu(s)| = 1$, and $\mu(s) = s$ if $\mu(s) \notin C$, (ii) $\mu(c) \subseteq S$ for every $c \in C$, and (iii) $\mu(s) = c$ if and only if $s \in \mu(c)$. That is, a matching simply specifies the college where each student is assigned or if the student is unmatched, and the set of students assigned to each college if any.

We say a matching μ is **blocked** by a pair of student s and college c if s strictly prefers c to $\mu(s)$ and either (i) c strictly prefers s to some $s' \in \mu(c)$ or (ii) $|\mu(c)| < q_c$ and s is acceptable to c. In words, the student s in the pair prefers college c over her assignment in μ, and college c prefers s either because it has a vacant seats or s is more preferred than another student assigned to c under μ. A matching μ is **individually rational** if for each student $s \in S$, $\mu(s) P_s \emptyset$ or $\mu(s) = \emptyset$ and for each $c \in C$, (i) $|\mu(c)| \leq q_c$ and (ii) $s \succ_c \emptyset$ for every $s \in \mu(c)$. A matching μ is **stable** if it is individually rational and is not blocked. A mechanism is a systematic way of assigning students to colleges. A stable mechanism is a mechanism that yields a stable matching with respect to reported preferences for every market.

We consider the **student-optimal stable mechanism (SOSM)**, denoted by ϕ, which is analyzed by Gale and Shapley (1962).[12]

- Step 1: Each student applies to her first choice college. Each college rejects the lowest-ranking students in excess of its capacity and all unacceptable students among those who applied to it, keeping the rest of the students temporarily (so students not rejected at this step may be rejected in later steps).

In general,

- Step t: Each student who was rejected in Step (t-1) applies to her next highest choice (if any). Each college considers these students *and* students who are temporarily held from the previous step together, and rejects the lowest-ranking students in excess of its capacity and all unacceptable students, keeping the rest of the students temporarily (so students not rejected at this step may be rejected in later steps).

[11] Note that given a responsive preference with quota $\succ_c$, we can always find a utility function $u_c : S \to \mathbb{R}$ with the property that for all $S', S'' \subseteq S$ such that $|S'|, |S''| \leq q_c$, $S' \succ_c S''$ if and only if $\sum_{s \in S'} u_c(s) > \sum_{s \in S''} u_c(s)$. We will use a particular additive representation of responsive preferences in Section 3.

[12] The SOSM is known to produce a stable matching that is unanimously most preferred by every student among all stable matchings (Gale and Shapley (1962)).

The algorithm terminates either when every student is matched to a college or every unmatched student has been rejected by every acceptable college. The algorithm always terminates in a finite number of steps. Gale and Shapley (1962) show that the resulting matching is stable. For two preference relations that induce the same pair of preference lists and quotas, the outcome of the algorithm is the same. Thus we sometimes write the resulting matching by $\phi(S, C, P, q)$. We denote by $\phi(S, C, P, q)(i)$ the assignment given to $i \in S \cup C$ under matching $\phi(S, C, P, q)$.

1.2 Manipulating the student-optimal stable mechanism

We illustrate two ways that the SOSM can be manipulated through a simple example.

Example 1. Consider the following market with two colleges c_1 and c_2, and five students $s_1, ..., s_5$. Suppose $q_{c_1} = 3$ and $q_{c_2} = 1$, and the preference lists of colleges are

$$P_{c_1} : s_1, s_2, s_3, s_4, s_5,$$
$$P_{c_2} : s_3, s_2, s_1, s_4, s_5,$$

and student preferences are

$$P_{s_1} : c_2, c_1,$$
$$P_{s_2} : c_1, c_2,$$
$$P_{s_3} : c_1, c_2,$$
$$P_{s_4} : c_1, c_2,$$
$$P_{s_5} : c_1, c_2.$$

The matching produced by the SOSM is

$$\mu = \begin{pmatrix} s_2 & s_3 & s_4 & s_1 \\ c_1 & c_1 & c_1 & c_2 \end{pmatrix},$$

which means that c_1 is matched to s_2, s_3 and s_4, c_2 is matched to s_1 and s_5 is unmatched.[13]

The first type of manipulation we will focus on is a **preference manipulation (manipulation via preference lists)**, first identified by Dubins and Freedman (1981). Suppose college c_1 submitted the following preference list:

$$P'_{c_1} : s_1, s_2, s_5, s_4, s_3,$$

[13]Similar notation is used throughout the paper.

while reporting its true quota, 3. Then the SOSM produces the following matching:

$$\mu' = \begin{pmatrix} s_1 & s_2 & s_5 & s_3 \\ c_1 & c_1 & c_1 & c_2 \end{pmatrix}.$$

If c_1 prefers $\{s_1, s_2, s_5\}$ to $\{s_2, s_3, s_4\}$, say because student s_1 is particularly desirable, then it can benefit by misreporting her preferences.[14]

Another way a stable mechanism can be manipulated, identified by Sönmez (1997a), is by **capacity manipulation (manipulation via capacities)**. Suppose that college c_1 states its quota as $q'_{c_1} = 1$, while reporting its preference list P_{c_1} truthfully. Then the SOSM produces the following matching:

$$\mu'' = \begin{pmatrix} s_1 & s_3 \\ c_1 & c_2 \end{pmatrix}.$$

If s_1 is more desirable than $\{s_2, s_3, s_4\}$, then college c_1 can benefit by reporting $q'_{c_1} = 1$.[15]

These two cases illustrate how the SOSM can be manipulated. We will consider both preference and capacity manipulations and their combination, and refer to this as a manipulation:

Definition 1. A college c can **manipulate the SOSM** if there exists a pair of a preference list and a quota (P'_c, q'_c) with $q'_c \in \{1, \ldots, q_c\}$ such that

$$\phi(S, C, (P'_c, P_{-c}), (q'_c, q_{-c}))(c) \succ_c \phi(S, C, P, q)(c).$$

We assume $q'_c \leq q_c$ because it is easily seen that reporting a quota larger than its true quota is never profitable.

1.3 Dropping strategies

The previous section presented an example illustrating preference and capacity manipulations. To study how likely it is that manipulations are successful, one way to begin may be to consider

[14]Note that relation $\{s_1, s_2, s_5\} \succ_{c_1} \{s_2, s_3, s_4\}$ is consistent with the assumption that $\succ_c$ is responsive.

[15]We note that, in this example, capacity manipulation may not benefit c_1 for some responsive preferences consistent with a fixed preference list P_{c_1}. Both $s_1 \succ_{c_1} \{s_2, s_3, s_4\}$ and $\{s_2, s_3, s_4\} \succ_{c_1} s_1$ are consistent with P_{c_1}, and capacity manipulation q'_{c_1} benefits c_1 in the former but not in the latter. Konishi and Ünver (2006) and Kojima (2007b) characterize the subclass of responsive preferences under which capacity manipulations benefit manipulationg colleges under stable mechanisms. For preference manipulations, on the contrary, there may exist manipulations that benefit the manipulating college for *all* responsive preferences that are consistent with a given preference list. In the current example, for instance, if c_1 declares $\tilde{P}_{c_1} : s_1, s_2, s_4, s_5$, then c_1 is matched to $\{s_1, s_2, s_4\}$. This is unambiguously preferred by c_1 to the match under truthtelling $\{s_2, s_3, s_4\}$ for any responsive preferences consistent with P_{c_1}.

all possible strategies of a particular college. For a college with a preference list of 5 students and a quota of 3 as in Example 1, this would involve considering all possible combinations of preference lists and quotas, which is 975 possible strategies, and verifying whether any of these strategies benefits the college.[16]

Fortunately, there is a general property of the manipulations we have discussed that allows us to greatly simplify the analysis and is one of the main building blocks of the analysis that follows. A reported pair of a preference list and a quota is said to be a **dropping strategy** if it simply declares some students who are acceptable under the true preference list as unacceptable. In particular, it does not misreport quotas or change the relative ordering of acceptable students or declare unacceptable students as acceptable.[17]

Returning to our example before, the outcome of the previous preference manipulation P'_{c_1} can be achieved by the dropping strategy $P''_{c_1} : s_1, s_2, s_5$, which simply drops students s_3 and s_4 from the true preference list. In the case of the previous capacity manipulation, if the college just dropped all students except s_1, then it would receive the same outcome as reducing the quota to 1. Note that in both of these cases, the original ordering of students is unchanged and only students on the original preference list are dropped.

The above observation turns out to be general for any stable mechanism. The following lemma which is formally stated and proved in Appendix B.1.1 shows that any manipulation can be replicated or improved upon by some dropping strategy.

Lemma 1 (Dropping strategies are exhaustive)**.** *Consider an arbitrary stable mechanism. Fix preferences of colleges other than c. Suppose the mechanism produces μ under some report of c. Then there exists a dropping strategy that produces a matching that c weakly prefers to μ according to its true preferences.*

Some intuition for this result can be seen in the above example. Consider a preference manipulation P'_{c_1}, which demoted s_3 to the bottom of the list of acceptable students. At one step of the SOSM, s_3 is rejected by P'_{c_1} because three other students applied to c_1 and s_3 is declared to be the least desirable. Then s_3 applies to c_2, which rejects s_1. Then s_1 applies to c_1 which accepts her, which benefits c_1. A similar chain of rejection and acceptances can be initiated if c_1 just declares s_3 as unacceptable. Similarly, additional chains of rejections and acceptances produced by underreporting of quotas can be replicated by declaring some students as unacceptable instead.

There are a few remarks about this lemma. First, while most of our analysis focuses on the SOSM, the lemma identifies a general property of any stable matching mechanism, which may be of independent interest. Second, a truncation strategy which only drops students from

[16]All possible combinations are calculated by observing that there are $5! + 5 \cdot 4 \cdot 3 \cdot 2 + 5 \cdot 4 \cdot 3 + 5 \cdot 4 + 5$ ways to submit preference lists and 3 ways to report capacities.

[17]Let (P_c, q_c) be a pair of the true preference list and true quota of college c. Formally, a dropping strategy is a report (P'_c, q_c) such that (i) sP_cs' and $sP'_c\emptyset$ imply sP'_cs', and (ii) $\emptyset P_cs$ implies $\emptyset P'_cs$.

the end of its preference list can replicate any profitable manipulation in a simpler one-to-one matching market (Roth and Vande Vate (1991)), and our lemma offers its counterpart in a more general many-to-one matching market.[18] Finally, and most importantly in the context of our work, the lemma allows us to focus on a particular class of strategies to investigate manipulations. In the previous example, we only need to consider $2^{|S|} = 2^5 = 32$ possible dropping strategies instead of all 975 strategies. In addition to being small in number, the class of dropping strategies turns out to be conceptually simple and analytically tractable.

2 Large markets

2.1 Regular markets

We have seen that a finite many-to-one matching market can be manipulated. To investigate how likely a college can manipulate the SOSM in large markets, we introduce the following random environment. A **random market** is a tuple $\tilde{\Gamma} = (C, S, \succ_C, k, \mathcal{D})$, where k is a positive integer and $\mathcal{D} = (p_c)_{c \in C}$ is a probability distribution on C. We assume that $p_c > 0$ for each $c \in C$.[19] Each random market induces a market by randomly generating preferences of each student s as follows (Immorlica and Mahdian (2005)):

- Step 1: Select a college independently from distribution $\mathcal{D}$. List this college as the top ranked college of student s.

In general,

- Step $t \leq k$: Select a college independently from distribution $\mathcal{D}$ until a college is drawn that has not been previously drawn in steps 1 through $t-1$. List this college as the t^{th} most preferred college of student s.

Student s finds these k colleges acceptable, and all other colleges unacceptable. For example, if $\mathcal{D}$ is the uniform distribution on C, then the preference list is drawn from the uniform distribution over the set of all preference lists of length k. For each realization of student preferences, a market with perfect information is obtained.

In the main text of the paper, we focus on the above procedure for distribution $\mathcal{D}$ to generate preferences of students for the sake of simplicity. For readers interested in how this

[18] Truncation strategies have found use in subsequent work, such as Roth and Peranson (1999), Roth and Rothblum (1999), Ma (2002), and Ehlers (2004). Truncation strategies may not be exhaustive when colleges may have a quota larger than one. For instance, in Example 1 we have seen that c_1 can be matched with $\{s_1, s_2, s_5\}$ by a dropping strategy P''_{c_1}, but such a matching or anything better for c_1 cannot be attained by a truncation strategy.

[19] We impose this assumption to compare our analysis with existing literature. Our analysis remains unchanged when one allows for probabilities to be zero.

can be generalized, we refer to the discussion in Section 4 and Appendix A.5, where we describe additional results under weaker assumptions.

A **sequence of random markets** is denoted by $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$, where $\tilde{\Gamma}^n = (C^n, S^n, \succ_{C^n}, k^n, \mathcal{D}^n)$ is a random market in which $|C^n| = n$ is the number of colleges.[20] Consider the following regularity conditions.

Definition 2. A sequence of random markets $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is **regular** if there exist positive integers k and $\bar{q}$ such that

(1) $k^n = k$ for all n,

(2) $q_c \leq \bar{q}$ for $c \in C^n$ and all n,

(3) $|S^n| \leq \bar{q}n$ for all n,[21] and

(4) for all n and $c \in C^n$, every $s \in S^n$ is acceptable to c.

Condition (1) assumes that the length of students' preference lists does not grow when the number of market participants grow. Condition (2) requires that the number of positions of each college is bounded across colleges and markets. Condition (3) requires that the number of students does not grow much faster than that of colleges. Condition (4) requires colleges to find any student acceptable, but preferences are otherwise arbitrary.[22] This paper focuses on regular sequences of random markets and makes use of each condition in our arguments. We will discuss directions in which these conditions can be weakened in the last section.

2.2 Main result

Consider the expected number of colleges that can manipulate the SOSM when others are truthful. Formally, such a number is defined as

$$\alpha(n) = E[\#\{c \in C | \phi(S, C, (P'_c, P_{-c}), (q'_c, q_{-c}))(c) \succ_c \phi(S, C, P, q)(c) \text{ for some } (P'_c, q'_c) \text{ in the induced market}\} | \tilde{\Gamma}^n].$$

The expectation is taken with respect to random student preferences, given the random market $\tilde{\Gamma}^n$. Note that we consider the possibility of manipulations under complete information,

[20] Unless specified otherwise, our convention is that superscripts are used for the number of colleges present in the market whereas subscripts are used for agents.

[21] As mentioned later, this condition can be relaxed to state: there exists $\tilde{q}$ such that $|S^n| \leq \tilde{q}n$ for any n (in other words, here we are assuming $\bar{q} = \tilde{q}$ just for expositional simplicity). In particular, the model allows situations in which there are more students than the total number of available positions in colleges.

[22] Condition (4) ensures that our economy has a large number of individually rational matchings, so potentially there is nontrivial scope for manipulations. It is possible to weaken this condition such that many, but not all colleges find all students to be acceptable.

that is, we investigate for colleges' incentives to manipulate when they know the preferences of every agent. The randomness of student preferences is used only to assess the frequency of situations in which colleges have incentives to manipulate.

Theorem 1. *Suppose that the sequence of random markets is regular. Then the expected proportion of colleges that can manipulate the SOSM when others are truthful, $\alpha(n)/n$, goes to zero as the number of colleges goes to infinity.*

This theorem suggests that manipulation of any sort within the matching mechanism becomes unprofitable to most colleges, as the number of participating colleges becomes large. In the Appendix A.1, we show that a manipulation that involves unraveling outside the centralized mechanism, named manipulation via pre-arranged matches (Sönmez (1999)), also becomes unprofitable as the market becomes large.

A limitation of Theorem 1 is that it does not imply that most colleges report true preferences *in equilibrium*. Equilibrium analysis is conducted in the next section.

The theorem has an implication about the structure of the set of stable matchings. It is well-known that if a college does not have incentives to manipulate the SOSM, then it is matched to the same set of students in all stable matchings. Therefore the following is an immediate corollary of Theorem 1.

Corollary 1. *Suppose that the sequence of random markets is regular. Then the expected proportion of colleges that are matched to the same set of students in all stable matchings goes to one as the number of colleges goes to infinity.*

Corollary 1 is referred to as a "core convergence" result by Roth and Peranson (1999). The main theorem of Immorlica and Mahdian (2005) shows Corollary 1 for one-to-one matching, in which each college has a quota of one.

The formal proof of Theorem 1 is in Appendix B.1. For the main text, we give an outline of the argument. We begin the proof by recalling that if a college can manipulate the SOSM, then it can do so by a dropping strategy. Therefore, when considering manipulations, we can restrict attention to a particular class of strategies which simply reject students who are acceptable under the true preferences.

The next step involves determining the outcome of dropping strategies. One approach might be to consider all possible dropping strategies and determine which ones are profitable for a college. While the set of dropping strategies is smaller than the set of all possible manipulations, this task is still daunting because the number of possible dropping strategies is large when there are a large number of students.

Thus we consider an alternative approach. We start with the student-optimal stable matching under the true preferences, and examine whether a college might benefit by dropping some student assigned in the student-optimal matching. Specifically, we consider a process where beginning with the student-optimal matching under true preferences, we drop some students from a particular college's preference list and continue the SOSM procedure starting with the original students rejected by the college because of the dropping. We refer to this process as **rejection chains.**[23]

For instance, in Example 1, starting with the student-optimal stable matching, if college c_1 dropped students $\{s_3, s_4\}$, let us first consider what happens to the least preferred student in this set, student s_4. If we examine the continuation of the SOSM, when s_4 is rejected, this student will propose to c_2, but c_2 is assigned to s_1 whom it prefers, so s_4 will be left unassigned. Thus, college c_1 does not benefit by rejecting student s_4 in this process, because this rejection does not spur another more preferred student to propose to c_1. On the other hand, consider what happens when college c_1 rejects student s_3. This student will propose to college c_2, who will reject s_1, freeing s_1 up to propose to c_1. By dropping s_3, college c_1 has created a new proposal which did not take place in the original SOSM procedure.

Our rejection chains procedure results in two cases: (1) a rejection never leads to a new proposal at the manipulating college as in the case of s_4, and (2) a rejection leads to a new student proposal at the college as in the case of s_3. The above example suggests that case (1) is never beneficial for the manipulating college, whereas case (2) may benefit the college. In the Lemma 3 in the Appendix, we make this intuition precise: if the rejection chain process does not lead to a new proposal (case (1)), then a college cannot benefit from a dropping strategy.[24] This result allows us to link the dropping strategy to the rejection chains algorithm. This connection gives us traction in large markets, as the number of cases to consider is bounded by a constant $2^{\overline{q}}$ for rejection chains, whereas there are $2^{|S^n|}$ potential dropping strategies in a market with n colleges, and the latter number may approach infinity as n approaches infinity.

The last step of the proof establishes that the probability that a rejection chain returns to the manipulating college is small when the market is large. To see the intuition for this step, suppose that there are a large number of colleges in the market. Then there are also a large number of colleges with vacant positions with high probability. We say that a college is popular if it is given a high probability in the distribution from which students preferences are drawn. Any student is much more likely to apply to one of those colleges with vacant positions rather than the manipulating college unless it is extremely popular in a large market, since there are a large number of colleges with vacant positions. Since every student is acceptable to any college by assumption, the rejection chains algorithm terminates without returning to the manipulating college if such an application happens. Thus, the probability that the algorithm returns to the manipulating college is very small unless the college is one of the small fraction of very popular colleges. Note that the expected proportion of colleges that can manipulate is equal to the sum of probabilities that individual colleges can manipulate. Together with our earlier reasoning, we conclude that the expected proportion of colleges that can successfully

[23]This process is formally defined in the Appendix B.1.2.

[24]We note that the manipulating college may not be made better off even when the rejection chain leads to a new proposal to the college (Case (2)).

manipulate converges to zero when the number of colleges grows.

Both Roth and Peranson (1999) and Immorlica and Mahdian (2005) attribute the lack of manipulability to the "core-convergence" property as stated in Corollary 1. While this interpretation is valid in one-to-one markets, the fact that a college is matched to the same set of students in all stable matchings is necessary but not sufficient for the college to have incentives for truth-telling in many-to-one markets. Instead our arguments show that lack of manipulability comes from the "vanishing market power" in the sense that the impact of strategically rejecting a student will be absorbed elsewhere and rarely affects the college that manipulated when the market is large.

Roth and Peranson (1999) analyze the NRMP data and argue that of the 3,000-4,000 participating programs, less than one percent could benefit by truncating preference lists or via capacities, assuming the data are true preferences. They also conduct simulations using randomly generated data in one-to-one matching, and observe that the proportion of colleges that can successfully manipulate quickly approaches zero as n becomes large. The first theoretical account of this observation is given by Immorlica and Mahdian (2005), who show Corollary 1 for one-to-one matching. Theorem 1 improves upon their results and fully explains observations of Roth and Peranson (1999) in the following senses: (1) it studies manipulations via preference lists in many-to-one markets, and (2) it studies manipulations via capacities. Neither of these points is previously investigated theoretically. Furthermore we strengthen assertions of Roth and Peranson (1999) and Immorlica and Mahdian (2005) by showing that large markets are immune to *arbitrary* manipulations and not just misreporting preference lists or misreporting capacities.

3 Equilibrium analysis

In the last section, we investigated individual colleges' incentives to manipulate the SOSM when all agents are truth-telling. As noted in the previous section, Theorem 1 does not necessarily mean that agents report true preferences *in equilibrium*. We now allow all participants to behave strategically and investigate equilibrium behavior in large markets. This section focuses on the simplest case to highlight the analysis of equilibrium behavior. The Appendix A.5 presents the more general treatment incorporating heterogeneous distributions of student preferences.

To investigate equilibrium behavior, we first define a normal-form game as follows. Assume that each college $c \in C$ has an additive utility function $u_c : 2^S \to \mathbb{R}$ on the set of subsets of students. More specifically, we assume that

$$u_c(S') \begin{cases} = \sum_{s \in S'} u_c(s) \text{ if } |S'| \leq q_c, \\ < 0 \text{ otherwise,} \end{cases}$$

where $u_c(s) = u_c(\{s\})$. We assume that $sP_cs' \iff u_c(s) > u_c(s')$. If s is acceptable to c, $u_c(s) > 0$. If s is unacceptable, $u_c(s) < 0$. Further suppose that utilities are bounded. Formally, $\sup u_c(s)$ is finite where the supremum is over the size of the market n, students $s \in S^n$, and colleges $c \in C^n$.

The normal-form game is specified by a random market $\tilde{\Gamma}$ coupled with utility functions of colleges and is defined as follows. The set of players is C, with von Neumann-Morgenstern expected utility functions induced by the above utility functions. All the colleges move simultaneously. Each college submits a preference list and quota pair. After the preference profile is submitted, random preferences of students are realized according to the given distribution $\mathcal{D}$. The outcome is the assignment resulting from ϕ under reported preferences of colleges and realized students preferences. We assume that college preferences and distributions of student random preferences are common knowledge, but colleges do not know realizations of student preferences when they submit their preferences.[25] Note that we assume that students are passive players and always submit their preferences truthfully. A justification for this assumption is that truthful reporting is weakly dominant for students under ϕ (Dubins and Freedman (1981), Roth (1982)).

First, let us define some additional notation. Let $P_C = (P_c)_{c\in C}$ and $P_{C-c} = (P_{c'})_{c'\in C, c'\neq c}$. Given $\varepsilon > 0$, a profile of preferences $(P^*_C, q^*_C) = (P^*_c, q^*_c)_{c\in C}$ is an ε-**Nash equilibrium** if there is no $c \in C$ and (P'_c, q'_c) such that

$$E\left[u_c(\phi(S, C, (P_S, P'_c, P^*_{C-c}), (q'_c, q^*_{-c}))(c))\right] > E\left[u_c(\phi(S, C, (P_S, P^*_C), q^*)(c))\right] + \varepsilon,$$

where the expectation is taken with respect to random preference lists of students.

Is truthful reporting an approximate equilibrium in a large market for an arbitrary regular sequence of random markets? The answer is negative, as shown by the following example.[26]

Example 2. Consider the following market $\tilde{\Gamma}_n$ for any n. There are n colleges and students. Preference lists of c_1 and c_2 are given as follows:[27]

$$P_{c_1} : s_2, s_1, \ldots,$$
$$P_{c_2} : s_1, s_2, \ldots.$$

Suppose that $p^n_{c_1} = p^n_{c_2} = 1/3$ and $p^n_c = 1/(3(n-2))$ for any $n \geq 3$ and each $c \neq c_1, c_2$.[28] With probability $[p^n_{c_1}p^n_{c_2}/(1-p^n_{c_1})] \times [p^n_{c_1}p^n_{c_2}/(1-p^n_{c_2})] = 1/36$, preferences of s_1 and s_2 are given

[25] Consider a game with incomplete information, in which each college knows other colleges' preferences only probabilistically. The analysis can be easily modified for this environment. See Appendix A.4.

[26] This example shows that Corollaries 3.1 and 3.3 in Immorlica and Mahdian (2005) are not correct.

[27] "..." in a preference list means that the rest of the preference list is arbitrary after those written explicitly.

[28] We use superscript n on p_c to indicate the probability of college c for market $\tilde{\Gamma}^n$. This notation is used in subsequent parts of the text.

by

$$P_{s_1} : c_1, c_2, \ldots,$$
$$P_{s_2} : c_2, c_1, \ldots.$$

Under the student-optimal matching μ, we have $\mu(c_1) = s_1$ and $\mu(c_2) = s_2$. Now, suppose that c_1 submits preference list $P'_{c_1} : s_2$. Then, under the new matching μ', c_1 is matched to $\mu'(c_1) = s_2$, which is preferred to $\mu(c_1) = s_1$. Since the probability of preference profiles where this occurs is $1/36 > 0$, regardless of $n \geq 3$, the opportunity for preference manipulation for c_1 does not vanish even when n becomes large. Therefore truth-telling is not an ε-Nash equilibrium if $\varepsilon > 0$ is sufficiently small, as c_1 has an incentive to deviate.

The above example shows that, while the *proportion* of colleges that can manipulate via preference lists becomes small, *for an individual college* the opportunity for such manipulation may remain large. Note on the other hand that this is consistent with Theorem 1 since the scope for manipulation becomes small for any $c \neq c_1, c_2$.[29]

A natural question is under what conditions we can expect truthful play to be an ε-Nash equilibrium. One feature of Example 2 is that c_1 and c_2 are popular and remain so even as the market becomes large. This suggests that it is the colleges which are extremely popular and remain so in a large market that may be able to manipulate.

Consider a market where there is not too much concentration of popularity among a small set of colleges. This feature of the market will reduce the influence of the exceptionally popular colleges that we observed in Example 2. One way to formally define this situation is to consider a sequence of random markets with the following property: there exists a finite bound T and fraction of a colleges such that

$$p_1^n / p_{[an]}^n \leq T, \tag{1}$$

for large n, where p_1^n is the popularity of the most popular college and $p_{[an]}^n$ is the popularity of the a^{th}-quantile college in market $\tilde{\Gamma}^n$. The condition means that the ratio of the popularity of the most popular college to the popularity of the college at the a^{th}-quantile does not grow without bound as the size of the market grows. This condition is satisfied if there are not a small number of colleges which are much more popular than all of the other colleges.

In Example 2, when n is large, $p_{c_1}^n / p_{c_2}^n = 1$, but $p_{c_1}^n / p_c^n = n - 2$ when c is not college c_1 or c_2, and this ratio grows without bound as $n \to \infty$. We will show later that inequality (1) is sufficient for truth-telling to be an ε-Nash equilibrium.

While this condition is easy to explain, it is not the most general condition that will ensure truthful play. Appendix A.3 presents additional examples with weaker assumptions. We can

[29] Capacity manipulation may remain profitable for some colleges in a large market as well. See the Appendix A.2 for an example illustrating this point.

describe a more general condition with the help of additional notation. Let

$$V_T(n) = \{c \in C^n | \max_{c' \in C^n} \{p^n_{c'}\}/p^n_c \leq T, |\{s \in S^n | cP_s s\}| < q_c\}.$$

In words, $V_T(n)$ is a random set which denotes the set of colleges sufficiently popular ex ante ($\max_{c'}\{p^n_{c'}\}/p^n_c \leq T$) but there are less potential applicants than the number of positions ($|\{s \in S^n | cP_s s\}| < q_c$) ex post.

Definition 3. A sequence of random markets is **sufficiently thick** if there exists $T \in \mathbb{R}$ such that

$$E[|V_T(n)|] \to \infty,$$

as $n \to \infty$.

The condition requires that the expected number of colleges that are desirable enough, yet have fewer potential applicants than seats, grows fast enough as the market becomes large. Consider a disruption of the market in which a student becomes unmatched. If the market is thick, such a student is likely to find a seat in another college that has room for her. Thus the condition would imply that a small disruption of the market is likely to be absorbed by vacant seats.[30] It is easy to verify that the market is not sufficiently thick in the example above. However, many types of markets satisfy the condition of sufficient thickness. For instance, if all student preferences are drawn from the uniform distribution, the market will be sufficiently thick. This is the environment first analyzed by Roth and Peranson (1999). The sufficient thickness condition is satisfied in more general cases which are described in Appendix A.3.

Theorem 2. *Suppose that the sequence of random markets is regular and sufficiently thick. Then for any $\varepsilon > 0$, there exists n_0 such that truth-telling by every college is an ε-Nash equilibrium for any market in the sequence with more than n_0 colleges.*

The proof of Theorem 2 (shown in Appendix B.3) is similar to that of Theorem 1 except for one point. With sufficient thickness, we can make sure that the rejection chain fails to return to any college with high probability, as opposed to only unpopular ones. This is because, with sufficient thickness, in a large market there are many vacant positions that are popular enough for students to apply to with a high probability and hence terminate the algorithm. The key difference is that in this section, we have an upper bound of the probability of successful manipulation for *every* college, while in the previous section, we have an upper bound only for unpopular colleges.

[30] This condition refers to the limit as the size of the market becomes large, so this notion is not relevant to a particular finite market. In particular, thickness and the size of the market are not related. It is even possible that the market does not become "thick" even when the market becomes large, in the sense that the limit in the above definition is finite as n goes to infinity.

We pursue generalizations of this result in the Appendix. First, we consider an incomplete information environment and show the conditions needed for truth-telling to be an ε-Bayesian equilibrium of the game (Appendix A.4). Second, we obtain a similar result when we allow for the possibility of pre-arrangement (Appendix A.6). Third, when we allow a coalition of colleges to manipulate, we extend the result to show the conditions under which coalitions of colleges will have little incentive to manipulate (Appendix A.7). The precise statements and details of each of these extensions is in the corresponding sections of the Appendix.

4 Discussion and Conclusion

Why do many stable matching mechanisms work in practice even though existing theory suggests that they can be manipulated in many ways? This paper established that the fraction of participants who can profitably manipulate the student-optimal stable mechanism (SOSM) in a large two-sided matching market is small under some regularity conditions. We further showed that truthful reporting is an approximate equilibrium in large markets that are sufficiently thick. These results suggest that a stable matching will be realized under the SOSM in a large market. Since a stable matching is efficient in general, the results suggest that the SOSM achieves a high level of efficiency in a large market.[31] Taken together, these results provide a strong case for the SOSM as a market design in large markets.

Based on the results in this paper, when do we expect certain matching mechanisms to be hard to manipulate? We will highlight three crucial assumptions here. First, our result is a limit result, and requires that the market is regular and large.[32] As discussed below, in a finite economy, the possibility of manipulation cannot be excluded without placing restrictive assumptions on preferences. Our results suggest that larger markets employing the SOSM will be less prone to manipulation than smaller markets, holding everything else the same.

Second, we have required that the length of the student preference list is bounded. As mentioned in the introduction, the primary motivation for this assumption is empirical. In the NRMP, the length of applicant preference lists is typically less than or equal to 15. In NYC, almost 75% of students ranked less than the maximum of 12 schools in 2003-04 and there were over 500 programs to choose from.[33] In Boston public schools, which recently adopted the

[31] In a somewhat different model of matching with price competition, Bulow and Levin (2006) show that stable matching mechanisms in their environment achieve high efficiency when there are a large number of market participants.

[32] It is easy to relax some of our conditions on the definition of a regular market. For instance, our assumption on bounded preference lists could be generalized to state: there is a sequence $(k^n)_{n=1}^{\infty}$ with $\lim_{n\to\infty}(k^n/\ln(n)) = 0$ such that, for any n and any student in S^n, the length of her preference list for the student is at most k^n. Similarly, the condition on the number of students can be relaxed to state: there exists $\tilde{q}$ such that $|S^n| \leq \tilde{q}n$ for any n.

[33] In subsequent years, the fraction of students with preference lists of less than 12 schools ranged between

SOSM, more than 90% of students ranked 5 or fewer schools at elementary school during the first and second year of the new mechanism, and there are about 30 different elementary schools in each zone.[34] The conclusions of our results are known to fail if the assumption of bounded preference lists is not satisfied and instead students regard every college as acceptable (Knuth, Motwani and Pittel (1990)). Roth and Peranson (1999) conduct simulations on random data illustrating this point.

One reason why students do not submit long preference lists is that it may be costly for them to do so. For example, medical school students in the U.S. have to interview to be considered by residency programs, and financial and time constraints can limit the number of interviews. Likewise, in public school choice, to form preference lists students need to learn about the programs they may choose from and in many instances they may have to interview or audition for seats. Another reason may be that there is an exogenous restriction on the length of student preference lists as in NYC. However, there is an additional concern for manipulation under such exogenous constraints on preference lists, since students may not be able to report their true preferences under the restriction.[35] We submit as an open question how the large market argument may be affected when there is a restriction on the length of the preference list (see Haeringer and Klijn (2007) for analysis in this direction, though in the finite market setting.)

Finally, our analysis relies on the distributional assumptions on student preferences. For any given number of colleges in the market, with enough data on the distribution of student preferences and college preferences, our methods can be adapted to obtain the bound on the potential gain from deviation from truth-telling. However, to verify whether a market is sufficiently thick requires making assumptions on college popularity as the market grows.

The main text has focused on a particular process for generating student preferences. Some distributions are excluded by this specification. For example, we assume that preferences of students are independently drawn from one another, and they are drawn from an identical distribution. The iid assumption and the particular process we analyzed in the main text excludes, for example, cases where students in a particular region are more likely to prefer colleges in that region than students in other regions.

In the Appendix A.5, however, we analyze a more general model that allows students to belong to one of several groups and draw their preferences from different distributions across

70% and 78%.

[34]There are a total of 84 different elementary schools in the entire city. A small handful of these programs are citywide, and can be ranked by a student who lives in any zone. There is also a small fraction of students who can apply across zones because they live near the boundary of two zones.

[35]NYC is a two-sided matching market using the SOSM which has an explicit restriction on the number of choices that can be ranked. Note that, however, even in NYC, in the first year of data from the new system about 75% of students rank less than the maximum number of 12 schools in the main round, so our analysis may still hold approximately. The only other example we are aware of is Spanish college admissions studied by Romero-Medina (1998).

groups. We believe that the model with group-specific student preferences may be important for studying a number of real-world markets, for participants may often have systematically different preferences according to their residential location, academic achievement, and other characteristics. For instance, in NYC, more than 80% of applicants rank a program that is in the borough of their residence as their top choice, for all four years we have data. We show that truthful reporting is an approximate equilibrium in large markets that are sufficiently thick when this sort of heterogeneity across groups is present.

Our result is known to fail when there is some form of interdependence of student preferences.[36] Obtaining the weakest possible condition on the distribution of preferences for our result to hold is a challenging problem for future work.

Another direction for generalization is to consider weakening the assumption of responsive preferences, or considering models of matching with transfers. With general substitutable preferences, however, telling the truth may not be a dominant strategy for students. Hatfield and Milgrom (2005) have recently shown that substitutable preferences together with the "law of aggregate demand" is sufficient to restore the dominant strategy property for students. It would be interesting to investigate if a result similar to ours holds in this environment.

Large markets and matching market design

Our results have established the virtues of the student-optimal stable mechanism in a large market. This may serve as one criterion to support its use as a market design, since other mechanisms may not share the same properties in a large market. To see this point, consider the so-called **Boston mechanism** (Abdulkadiroğlu and Sönmez (2003)), which is often used for real-life matching markets. The mechanism proceeds as follows:[37]

- Step 1: Each student applies to her first choice college. Each college rejects the lowest-ranking students in excess of its capacity and all unacceptable students.

In general,

- Step t: Each student who was rejected in the last step proposes to her next highest choice. Each college considers these students, *only as long as* there are vacant positions not filled by students who are already matched by the previous steps, and rejects the lowest- ranking students in excess of its capacity and all unacceptable students.

[36]Immorlica and Mahdian (2005) present a model in which preferences cannot be generated in our procedure and the fraction of colleges that can manipulate does not go to zero even as the size of the market goes to infinity. Since their environment is a special case of ours, it shows that some assumptions similar to ours are needed to obtain our results.

[37]With slight abuse of terminology, we will refer to this mechanism where colleges rank students as the Boston mechanism even though the Boston mechanism was introduced as a one-sided matching mechanism, that is, college preferences are given as an exogenously given priority structure.

The algorithm terminates either when every student is matched to a college or when every unmatched student has been rejected by every acceptable college.[38]

Under the Boston mechanism, colleges have no incentives to manipulate either via preference lists or via capacity even in a small market with an arbitrary preference profile (Ergin and Sönmez (2006)). Nevertheless, we argue that this mechanism performs badly both in small and large markets. The problem is that students have incentives to misrepresent their preferences, and there is evidence that some participants react to these incentives (Abdulkadiroğlu, Pathak, Roth and Sönmez (2006)). An example in the Appendix A.8 shows that students' incentives to manipulate the Boston mechanism remain large even when the number of colleges increases in a regular sequence of markets.

We view this paper as one of the first attempts in the matching literature to show how performance of matching mechanisms in large economies can be used to compare mechanisms. One traditional approach in the matching literature is to find restrictions on preferences to ensure that a mechanism produces a desirable outcome. Alcalde and Barberà (1994) investigate this question focusing on preference manipulation, and Kesten (2006), Kojima (2007b), and Konishi and Ünver (2006) focus on capacity manipulation and manipulation via pre-arranged matches. One message from these papers is that the conditions that prevent the possibility of manipulation are often quite restrictive. In this paper, we have developed a different approach which uses large market arguments to obtain possibility results despite a series of impossibility results in the literature.

More generally, we think that the kind of large market arguments similar to the current paper will be fruitful in future work on matching and related allocation mechanisms. In the problem of allocating indivisible objects such as university housing, Kojima and Manea (2006) show that the probabilistic serial mechanism, which has desirable efficiency and fairness properties (Bogomolnaia and Moulin (2001)), also has a desirable form of incentive compatibility in a large market. In the kidney exchange problem (Roth, Sönmez and Ünver (2004)), Roth, Sönmez and Ünver (2007) show that efficiency can be achieved by conducting only kidney exchanges of small sizes when the number of incompatible patient-donor pairs is sufficiently large. Both of these are cases where large market arguments support a particular matching market design.

There are also open questions where large market analysis may yield new insights. Roth (2007) suggests that large market arguments may be useful to understand why a stable matching was always found in the NRMP even though the existence of couples can make the set of stable matchings empty.[39] In the school choice setting with indifferences, Erdil and Ergin (2008)

[38]Note the difference between this mechanism and the SOSM. At each step of the Boston mechanism, students who are not rejected are *guaranteed* positions; the matches of these students and colleges are permanent rather than temporary, unlike in the student-optimal stable mechanism.

[39]To ensure the existence of stable matchings with couples, we need restrictive assumptions on preferences of couples (Klaus and Klijn (2005)). Kojima (2007a) develops an algorithm to find stable matchings with couples without imposing assumptions on preferences which might be a first step to investigate this issue in

propose a new procedure to construct a student-optimal matching. While there are potentially large efficiency gains from their procedure (Abdulkadiroğlu et al. (2008)), in their mechanism, it is not a dominant strategy for students to reveal their preferences truthfully. Evaluating incentive properties of this new procedure in large markets may be an interesting direction of future research.

As market design tackles more complex environments, it will be harder to obtain finite market results on the properties of certain mechanisms. The current paper explored an alternative approach, which is based on a large market assumption. Since many markets of interest can be modeled as large markets, explicitly analyzing the limit properties will be a useful approach to guide policymakers and help evaluate designs in these environments.

large markets.

Appendix

Appendix A contains additional results referred to in the main text. Appendix B contains all proofs from the main text except the proof of Theorem 4 and Theorem 6.

We first introduce some notation. First, non-strict counterparts of P_s and $\succ_c$ are denoted by R_s and $\succeq_c$, respectively. For any pair of matchings μ and μ' and for any $c \in C$, we write $\mu \succ_c \mu'$ if and only if $\mu(c) \succ_c \mu'(c)$. Similarly, for any $s \in S$, we write $\mu P_s \mu'$ if and only if $\mu(s) P_s \mu'(s)$. Without loss of generality, we assume the set of colleges C are *ordered in decreasing popularity*: if $c' < c$, then $p_{c'} \geq p_c$. With abuse of notation, we write $c = m$, $c > m$ and $c < m$ for $m \in \mathbb{N}$ to mean, respectively, that c is the mth college, c is ordered after the mth college and c is ordered before the m college. We sometimes write p_m, which is the probability associated with the mth college.

A Appendix: Additional results

A.1 Manipulation via pre-arranged matches

When colleges seek more than one student, there is concern for manipulation not only within the matching mechanism, but also outside the formal process. Sönmez (1999) introduces the idea of manipulation via pre-arranged matches. Suppose that c and s arrange a match before the central matching mechanism is executed. Then s does not participate in the centralized matching mechanism and c participates in the centralized mechanism with the number of positions reduced by one. The SOSM is **manipulable via pre-arranged matches**, or **manipulable via pre-arrangement**, that is, for some market (S, C, P, q), college $c \in C$ and student $s \in S$ we have

$$\phi(S \setminus s, C, P_{-s}, (q_c - 1, q_{-c}))(c) \cup s \succ_c \phi(S, C, P, q)(c), \text{and}$$
$$cR_s\phi(S, C, P, q)(s).$$

In words, both parties that engage in pre-arrangement have incentives to do so: the student is at least as well off in pre-arrangement as when she is matched through the centralized mechanism, and the college strictly prefers s and the assignment of the centralized mechanism to those without pre-arrangement. Sönmez (1999) shows that any stable mechanism is manipulable via pre-arrangement.

In some markets, matching outside the centralized mechanism is discouraged or even legally prohibited. Even so, the student and college can effectively "pre-arrange" a match by listing each other on the top of their preference lists under stable mechanisms such as the SOSM. Thus the scope of manipulation via pre-arrangement is potentially large.

However, we have the following positive result in large markets.

Theorem 3. *Suppose that the sequence of random markets is regular. Then the expected proportion of colleges that can manipulate the SOSM via pre-arranged matches (when other colleges do not manipulate) goes to zero as the number of colleges goes to infinity.*

The intuition is similar to that of Theorem 1. It can be shown that any student involved in pre-arrangement under the SOSM is strictly less preferred by c to any student who would be matched in the absence of the pre-arrangement (Lemma 8). Therefore, in order to profitably manipulate, c should be matched to a better set of students in the central matching. By a similar reasoning to Theorem 1, the probability of being matched to better students in the centralized mechanism is small in a large market for most colleges.

A.2 Manipulation via capacities and pre-arranged matches without sufficient thickness

The next example shows that, when we do not have sufficient thickness, manipulations via capacities or pre-arrangement may be profitable for some colleges even in a large market.

Example 3. Consider the following market $\tilde{\Gamma}_n$ for any n. $|C^n| = |S^n| = n$. $q_{c_1} = 2$ and $q_c = 1$ for each $c \neq c_1$. c_1's preference list is

$$P_{c_1} : s_1, s_2, s_3, s_4, \ldots,$$

and $s_1 \succ_{c_1} \{s_2, s_3\}$ (and hence $\{s_1, s_4\} \succ_{c_1} \{s_2, s_3\}$).

c_2's preferences are

$$P_{c_2} : s_3, s_1, s_2, \ldots.$$

Further suppose that $p^n_{c_1} = p^n_{c_2} = 1/3$ and $p^n_c = 1/(3(n-2))$ for any n and each $c \neq c_1, c_2$.

With the above setup, with probability $[p^n_{c_1} p^n_{c_2}/(1-p^n_{c_2})] \times [p^n_{c_1} p^n_{c_2}/(1-p^n_{c_1})]^3 = 1/6^4$, students preferences are given by

$$P_{s_1} : c_2, c_1, \ldots,$$
$$P_{s_2} : c_1, c_2, \ldots,$$
$$P_{s_3} : c_1, c_2, \ldots,$$
$$P_{s_4} : c_1, c_2, \ldots.$$

If everyone is truthful, then c_1 is matched to $\{s_2, s_3\}$. Now

(1) Suppose that c_1 reports a quota of one. Then c_1 is matched to s_1, which is preferred to $\{s_2, s_3\}$.

(2) Suppose that c_1 pre-arranges a match with s_4. Then c_1 is matched to $\{s_1, s_4\}$, which is preferred to $\{s_2, s_3\}$.

Since the probability of preference profiles where this occurs is $1/6^4 > 0$ regardless of $n \geq 3$, the opportunity for manipulations via capacities or pre-arrangement for c_1 does not vanish when n becomes large.[40]

A.3 Examples of sufficiently thick markets

The following is a leading example of sufficient thickness.

Example 4 (Nonvanishing proportion of popular colleges)**.** The sequence of random markets is said to have **nonvanishing proportion of popular colleges** if there exists $T \in \mathbb{R}$ and $a \in (0, 1)$ such that for large n

$$p_1^n / p_{[an]}^n \leq T,$$

where $[x]$ denotes the largest integer that does not exceed x. This condition is satisfied if there are not a small number of colleges which are much more popular than all of the other colleges.

There are even sufficiently thick markets where the proportion of popular colleges converges to zero, provided that the convergence is sufficiently slow. We present one such example next.

Example 5. *Consider a sequence of random markets such that there exists $T \in \mathbb{R}$ such that for large n,*

$$p_1^n / p_{[\gamma n / \ln n]}^n \leq T$$

where $\gamma > 0$ is a sufficiently large constant.

Proposition 1. *The sequences of random markets in Examples 4 and 5 are sufficiently thick.*

The intuition for this proposition is the following. By assumption, there are a large number of ex ante popular colleges. With high probability, a substantial part of the positions of these colleges will be vacant. This makes the market thick by having a large number of vacant positions in fairly popular colleges in expectation.

A.4 Equilibrium analysis with incomplete information

The main text of the paper investigated ε-Nash equilibrium under complete information about college preferences. Our result also can be stated in terms of a Bayesian game, in which college preferences are private information.

[40]Manipulation via preference list is also possible in this example. Suppose c_1 reports preferences declaring s_2 and s_3 are unacceptable, such that

$$P'_{c_1} : s_1, s_4, \ldots .$$

Then c_1 is matched to $\{s_1, s_4\}$, which is preferred to $\{s_2, s_3\}$.

The Bayesian game is specified by $(C, S, (\mathcal{U}_c)_{c\in C}, F, k, \mathcal{D})$, where $\mathcal{U}_c$ is the set of possible utility functions for college c, and $F : \prod_{c\in C} \mathcal{U}_c \to [0, 1]$ is the distribution of utility types. We assume that each $u_c \in \mathcal{U}_c$ is an additive function $u_c : 2^S \to \mathbb{R}$ on the set of subsets of students. More specifically, we assume that

$$u_c(S') \begin{cases} = \sum_{s\in S'} u_c(s) \text{ if } |S'| \leq q_c, \\ < 0 \text{ otherwise,} \end{cases}$$

where $u_c(s) = u_c(\{s\})$. We assume that $sP_cs' \iff u_c(s) > u_c(s')$. If s is acceptable to c, $u_c(s) > 0$. If s is unacceptable, $u_c(s) < 0$. Further, we suppose that $\sup_{n\in\mathbb{N},s\in S^n,c\in C^n,u_c\in\mathcal{U}_c} u_c(s)$ is finite.

The set of players is C, with von Neumann-Morgenstern expected utility functions in $(\mathcal{U}_c)_{c\in C}$ drawn from distribution F. All the colleges move simultaneously. College c submits a preference list and quota pair (P'_c, q'_c) with $1 \leq q'_c \leq q_c$ after its realization of utility u_c, but without observing the utilities realized by the other colleges. A strategy for college c is a report $(P'_c(u_c), q'_c(u_c))$ for each possible utility function $u_c \in \mathcal{U}_c$.

After colleges submit a preference profile, random preferences of students are realized according to the given distribution $\mathcal{D}$. The outcome is the assignment resulting from ϕ under reported preferences of colleges and realized students preferences. We assume the distribution of college preferences is independent of the distribution of student preferences, and both distributions are common knowledge. Moreover, a college does not know realizations of student preferences. As in the main text, we assume that students are passive players and always submit their preferences truthfully.

Given $\varepsilon > 0$, a strategy profile $(P^*_c(u_c), q^*_c(u_c))_{c\in C,u_c\in\mathcal{U}_c}$ is an **ε-Bayes Nash equilibrium** if there is no $c \in C$, $u_c \in \mathcal{U}_c$ and (P'_c, q'_c) such that

$$\begin{aligned} &E[u_c(\phi(S, C, (P_S, P'_c, (P^*_{c'}(u_{c'}))_{c'\in C\setminus c}), (q'_c, (q^*_{c'}(u_{c'}))_{c'\in C\setminus c})))] \\ &\qquad\qquad > E[u_c(\phi(S, C, (P_S, (P^*_{c'}(u_{c'}), q^*_{c'}(u_{c'}))_{c'\in C})))] + \varepsilon, \end{aligned}$$

where the expectation is taken with respect to random preference lists of students and distribution F of college preferences.

We say that a strategy $(P'_c(u_c), q'_c(u_c))$ is truth-telling if the college reports the preferences (P_c, q_c) represented by utility function u_c, for each $u_c \in \mathcal{U}_c$. Now, we can restate Theorem 2 for the Bayesian game.

Theorem 4. *Suppose that the sequence of random markets is regular and sufficiently thick. Then for any $\varepsilon > 0$, there exists n_0 such that truth-telling by every college is an ε-Bayes Nash equilibrium for any market in the sequence with more than n_0 colleges.*

Proof. From the proof of Theorem 2, we know that for any $\varepsilon > 0$, there exists n_0 such that for each realization of the utilities of colleges, truth-telling is an ε-Nash equilibrium for any market in the sequence with more than n_0 colleges for that realization of colleges' utilities. Since the result holds for each realization of utilities and we can find n_0 uniformly across utility realizations, truth-telling by all colleges is an ε-Bayesian Nash as well. $\square$

A.5 Weakening distributional assumptions

In the main text, we have focused on a simple case in which student preferences are drawn from the same distribution. This section extends our analysis to cases in which student preferences are drawn from a number of different distributions.

The model is the same as before except for how student preferences are drawn. We now defined a random market as $\tilde{\Gamma}^n = (C^n, S^n, \succ_{C^n}, k^n, (\mathcal{D}^n(r))_{r=1}^{R^n})$, where R^n is a positive integer. Each random market is endowed with R^n different distributions. To represent student preferences, we partition students into R^n regions, where each student is a member of exactly one region.[41] Write $\mathcal{D}^n(r) = (p_c^n(r))_{c \in C^n}$ as the probability distribution on C^n for students in region r. For each student $s \in S^n$ in region r, we construct preferences of s over colleges as described below:

- Step 1: Select a college independently according to $\mathcal{D}^n(r)$. List this college as the top ranked college of student s.

In general,

- Step $t \leq k$: Select college independently according to $\mathcal{D}^n(r)$ until a college is drawn that has not been previously drawn in steps 1 through $t-1$. List this college as the tth most preferred college of student s.

We will refer to this method of generating student preferences as the **model with heterogenous student preference distributions**. The case with $R^n = 1$ for all n corresponds to our earlier model with one distribution for student preferences.

Our regularity assumptions extend naturally: in addition to conditions in the previous definition, we also require that, for some positive integer $\bar{R}$, $R^n = \bar{R}$ for every n in a regular market. Finally, our assumption of sufficient thickness generalizes easily to the current environment. Let

$$V_T(n) = \{c \in C^n | p_1^n(r)/p_c^n(r) \leq T \quad \text{for all } r, |\{s \in S^n | cP_s s\}| < q_c\},$$
$$Y_T(n) = |V_T(n)|.$$

[41] We frame the heterogeneity of student preferences in terms of multiple regions where students live. Of course alternative interpretations are possible, such as heterogeneity depending on medical specialties, gender, race or academic performance or combinations of these characteristics.

Definition 4. A sequence of random markets is **sufficiently thick** if there exists $T \in \mathbb{R}$ such that

$$E[Y_T(n)] \to \infty,$$

as $n \to \infty$.

Definition 4 is a multi-region generalization of sufficient thickness for one-region setting. The following examples satisfy this version of sufficient thickness.

Example 6 (Two regions with opposite popularity)**.** Fix an arbitrary quota q_c for each college c. There are two regions, $\bar{R} = 2$. $C^n = \{1, 2, \dots, n\}$ and the probability distributions are:

$$p^1_c(1) = \frac{n - c + 1}{\sum_{c' \in C^n}(n - c' + 1)} = \frac{n - c + 1}{\frac{n(n+1)}{2}},$$
$$p^1_c(2) = \frac{c}{\sum_{c' \in C^n} c'} = \frac{c}{\frac{n(n+1)}{2}}.$$

Students in the first region prefer the first college over the second college and so forth on average, while students in the second region have the opposite preferences. There is an extreme form of differences in preferences in this market.

Example 7 (Multiple regions with within-region symmetry)**.** Fix an arbitrary quota q_c for each college c. Assume there are $\bar{R}$ regions, $\bar{R} \geq 2$. Each college is based in one of the regions. Let $r(c)$ be the region in which college c is. Let $\tilde{p}_m(r)$, $r, m \in \{1, \dots, \bar{R}\}$ be strictly positive for every r, m. From this, we define the probability $p^n_c(r)$ for any $n \in \mathbb{N}$ as follows:

$$p^n_c(r) = \frac{\tilde{p}_{r(c)}(r)}{\sum_{m \in \bar{R}} \tilde{p}_m(r) \nu^n_m},$$

where $\nu^n_m = |\{c \in C^n | r(c) = m\}|$ denotes the number of colleges in $\tilde{\Gamma}^n$ that is based in region m.

This environment has the following interpretation. Each college is based in one of the regions, and each student lives in one region. Colleges in a given region are equivalent to one another. The "base popularity" of a college in region m for a student living in region r is given by $\tilde{p}_m(r)$. Then we normalize these to obtain $p^n_c(r)$ by the above equation. For any pair of colleges c and c' and region r, we have that

$$p^n_c(r)/p^n_{c'}(r) = \tilde{p}_{r(c)}(r)/\tilde{p}_{r(c')}(r).$$

Such heterogeneous preferences may be present in labor markets or in large urban school districts, where students in the same region have similar preferences while substantial differences are present across regions.

Proposition 2. *Sequences of random markets in Examples 6 and 7 are sufficiently thick.*

These are among the simplest examples incorporating heterogeneity. The two region case shows that directly opposing preferences satisfy sufficient thickness. The multiple region case illustrates that a great deal of heterogeneity in student preferences is allowed.

The equilibrium analysis in the one-region setting (Theorem 2) extends to heterogeneous preference distributions such that the market is sufficiently thick.

Theorem 5. *Consider the model where student preferences are from heterogeneous distributions. Suppose that the sequence of random markets is regular and sufficiently thick. Then for any $\varepsilon > 0$, there exists n_0 such that truth-telling is an ε-Nash equilibrium for any market with more than n_0 colleges.*

A.6 Pre-arrangement in a sufficiently thick market

This section states and proves a result similar in spirit to Theorem 2 for pre-arrangement.

Theorem 6. *Consider the model where student preferences are from heterogeneous distributions. Suppose that $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is regular and sufficiently thick. Consider the SOSM. For any $\varepsilon > 0$, there exists n_0 such that for any $n > n_0$ and $c \in C^n$, the probability that c can profitably manipulate via pre-arrangement is smaller than ε.*

Proof. In the proof of Theorem 3, we have shown that for each college c, the probability of successful manipulation is at most π_c. By Lemma 10 and sufficient thickness, there exists a T such that

$$\pi_c \leq \frac{4\,[T\bar{q}(2^{\bar{q}} - 1) + 1]}{E[Y_T(n)]}.$$

Given ε, the definition of sufficient thickness implies that $E[Y_T(n)] \to \infty$, so $\pi_c \to 0$. □

A.7 Manipulations by coalitions

The basic model shows that individual colleges have little opportunity to manipulate a large market. One natural question is whether coalitions of colleges can manipulate by coordinating their reports. Formally, a coalition $\bar{C} \subseteq C$ manipulates the market (S, C, P, q) if there exists $(P'_{\bar{C}}, q'_{\bar{C}}) = (P'_c, q'_c)_{c \in \bar{C}}$ such that

$$\phi(S, C, (P'_{\bar{C}}, P_{-\bar{C}}), (q'_{\bar{C}}, q_{-\bar{C}})) \succ_c \phi(S, C, P, q),$$

for some $c \in \bar{C}$.

The notion of coalitional manipulation we consider allows for a broad range of coalitions, for a coalition is said to manipulate even if only some of its members are made strictly better off and others in the coalition are made strictly worse off when they misreport their preferences jointly.

Theorem 7. *Consider the model where student preferences are from heterogeneous distributions. Suppose that the sequence of random markets is regular and sufficiently thick. Consider the SOSM. Then, for any positive integer m and any $\varepsilon > 0$, there exists n_0 such that for any $n > n_0$ and $\bar{C} \subseteq C^n$ with $|\bar{C}| \leq m$, the probability that $\bar{C}$ can profitably manipulate is smaller than ε.*

Our result shows that successful coalitional manipulation is rare: with high probability, not a single college in the coalition is made strictly better off. Thus it is hard for coalitions to manipulate even when monetary transfers are possible among colleges.

A.8 Manipulating the Boston mechanism in large markets

The following example shows that students have incentives to manipulate the Boston mechanism even in large markets.

Example 8. Consider market $\tilde{\Gamma}^n$, where $|S^n| = |C^n| = n$ for each n. $q_c^n = 1$ for every n and $c \in C^n$. Preference lists are common among colleges and given by

$$P_c : s_1, s_2, \ldots, s_n,$$

for every $c \in C^n$.

Let $p_{c_1}^n = (1/2)^{1/n}, p_{c_2}^n = (1 - (1/2)^{1/n})(1/2)^{1/n}$, and $p_c^n = (1 - p_{c_1}^n - p_{c_2}^n)/(n-2)$ for each $c \neq c_1, c_2$. Then, with probability $[p_{c_1}^n p_{c_2}^n/(1 - p_{c_1}^n)]^n = 1/4$, students preferences are

$$P_s : c_1, c_2, \ldots,$$

for each $s \in S^n$. If every student is truth-telling, then s_1 and s_2 are matched to c_1 and c_2, respectively, and other students are matched to their third or less preferred choices. If $s \neq s_1, s_2$ deviates from truth-telling unilaterally and reports preference list

$$P_s : c_2, \ldots,$$

then s is matched to her second choice c_2, which is preferred to the match under truth-telling. This occurs with probability of at least $1/4$, and every student except s_1 and s_2 has an incentive not to be truth-telling.

B Appendix: Proofs

B.1 Proof of Theorem 1

We prove Theorem 1 through several steps. Specifically, we prove three key lemmas, Lemmas 1, 3 and 7 and then use them to show the theorem. For the proof of Lemmas 1 and 3, we keep the set of students and colleges fixed, and refer to markets in terms of preference profiles only.

Let (P_c, q_c) be a pair of the true preference list and true quota of college c. A report (P'_c, q_c) is said to be a **dropping strategy** if (i) sP_cs' and $sP'_c\emptyset$ imply sP'_cs', and (ii) $\emptyset P_cs$ implies $\emptyset P'_cs$. A dropping strategy does not modify quotas and simply declares some students who are acceptable under P_c as unacceptable. It does not change the relative ordering of acceptable students or declare unacceptable students as acceptable.

B.1.1 Lemma 1: Dropping strategies are exhaustive

Given a stable mechanism φ, denote the matching under φ with respect to reported profile (P, q) by $\varphi(P, q)$. The formal statement of Lemma 1 in the main text is as follows.

Lemma 1. *Consider an arbitrary stable mechanism φ. Fix preference profile (P, q). For some report $(\tilde{P}_c, \tilde{q}_c)$ of c, suppose that $\varphi((\tilde{P}_c, P_{-c}), (\tilde{q}_c, q_{-c})) = \mu$. Then there exists a dropping strategy (P'_c, q_c) such that $\varphi((P'_c, P_{-c}), q) \succeq_c \mu$.*

Proof. Construct dropping strategy (P'_c, q_c) such that P'_c lists all the students it is matched to under μ who are acceptable ($\{s \in \mu(c)|sP_c\emptyset\}$) in the same relative order as in P_c, and reports every other student as unacceptable. Let $(P', q) = (P'_c, P_{-c}, q)$.

We will show $\varphi(P', q)(c)$ is equal to $\{s \in \mu(c)|sP_c\emptyset\}$. This equality implies that $\varphi(P', q) \succeq_c \mu$, since $\varphi(P', q)(c)$ can only differ from $\mu(c)$ in having no unacceptable students under the true preference list P_c. The proof proceeds in two steps.

First, consider the matching μ' obtained from μ by having c keep only the students in $\mu(c)$ that are acceptable under P_c. That is,

$$\mu'(c') = \begin{cases} \{s \in \mu(c)|sP_c\emptyset\} & c' = c, \\ \mu(c') & c' \neq c. \end{cases}$$

Consider the properties of μ' under (P', q). First, μ' is individually rational under (P', q). Second, there is no blocking pair involving c, since $\mu'(c)$ is exactly the set of students acceptable under P'_c. However, there may be a blocking pair involving another college and students who are unmatched. In this case, we can define a procedure which ultimately yields a matching that is stable under (P', q):[42]

[42] This procedure parallels Blum, Roth and Rothblum (1997)'s analogous vacancy-chain procedure for one-to-one matching markets and Cantala (2004)'s algorithm for many-to-one markets.

Starting from $\mu'_0 \equiv \mu'$, let S_0 be the set of students that can be part of a blocking pair in μ'_0. Pick some $s \in S_0$. Let college c' be the college that student s prefers the most within the set of colleges that can be part of a blocking pair with s. Construct μ'_1 by assigning student s to c' and if $|\mu'_0(c')| = q_{c'}$, then let c' reject its least preferred student. By construction, μ'_1 is individually rational, college c is matched to the same set of students, college c' strictly prefers $\mu'_1(c')$ over $\mu'_0(c')$ and the assignments of all other colleges are unchanged. As with μ'_0, the only blocking pair of μ'_1 involves unmatched students. If there are blocking pairs of μ'_1, then repeat the same procedure to construct μ'_2, and so on.

At each repetition, the new matching is individually rational, college c is matched to the same set of students, one other college strictly improves and the remaining colleges are not worse off since each blocking pair involves an unassigned student. Since colleges can strictly improve only a finite number of times, this procedure terminates in finite time. The ultimate matching μ'' is stable in (P', q) because it is individually rational and there are no blocking pairs. Moreover, since college c obtains the same matching in each repetition,

$$\mu''(c) = \mu'(c). \tag{2}$$

The second step of the proof utilizes the fact that for any college, the same number of students are matched to it across different stable matchings (Roth (1984a)). Since μ'' is stable in (P', q) and $\varphi(P', q)$ is stable in (P', q) by definition, $|\mu''(c)| = |\varphi(P', q)(c)|$. Since there are just $|\mu''(c)|$ acceptable students under P'_c, this implies that

$$\mu''(c) = \varphi(P', q)(c). \tag{3}$$

Equations (2) and (3) together imply that

$$\varphi(P', q)(c) = \mu'(c) = \{s \in \mu(c) | s P_c \emptyset\}.$$

As a result,

$$\varphi(P', q)(c) \succeq_c \mu(c),$$

since $\varphi(P', q)(c)$ can only differ from $\mu(c)$ in having no unacceptable students under the true preference list P_c.

$\square$

B.1.2 Lemma 3: Rejection chains

For preference profile (P, q), let μ be the student-optimal stable matching. Let B_c^1 be an arbitrary subset of $\mu(c)$. The **rejection chains** algorithm with input B_c^1 is defined as follows.

Algorithm 1. Rejection Chains

(1) Initialization:

(a) μ is the student-optimal stable matching, and B_c^1 is a subset of $\mu(c)$. Let $i = 0$. Let c reject all the students in B_c^1.

(2) Increment i by one.

(a) If $B_c^i = \emptyset$, then terminate the algorithm.

(b) If not, let s be the least preferred student by c among B_c^i, and let $B_c^{i+1} = B_c^i \setminus s$.

(c) Iterate the following steps (call this iteration "Round i").

i. Choosing the applied:

A. If s has already applied to every acceptable college, then finish the iteration and go back to the beginning of Step 2.

B. If not, let c' be the most preferred college of s among those which s has not yet applied while running the SOSM or previously within this algorithm. If $c' = c$, terminate the algorithm.

ii. Acceptance and/or rejection:

A. If c' prefers each of its current mates to s and there is no vacant position, then c' rejects s; go back to the beginning of Step 2c.

B. If c' has a vacant position or it prefers s to one of its current mates, then c' accepts s. Now if c' had no vacant position before accepting s, then c' rejects the least preferred student among those who were matched to c'. Let this rejected student be s and go back to the beginning of Step 2c. If c' had a vacant position, then finish the iteration and go back to the beginning of Step 2.

Algorithm 1 terminates either at Step 2a or at Step 2(c)iB. We say that Algorithm 1 **does not return to c** if it terminates at Step 2a and it **returns to c** if it terminates at Step 2(c)iB.

Lemma 2. *Consider college c and suppose that under (P, q), Algorithm 1 with every possible subset of $\mu(c)$ as an input fails to return to c. For any dropping strategy (P'_c, q_c), let $B_c^1 = \{s \in \mu(c) | \emptyset P'_c s\}$ be non-empty and let μ' be a matching obtained at the end of Algorithm 1 with input B_c^1. Then, under (P'_c, P_{-c}, q),*

(1) *μ' is individually rational,*

(2) *no $c' \neq c$ is a part of a blocking pair of μ', and*

(3) *if (s, c) blocks μ', then $[\arg\min_{P_c} \mu(c)] P_c s$,*[43] *and*

[43] For any binary relation R on X and $X' \subseteq X$, $\arg\min_R X' = \{x \in X' | yRx \text{ for any } y \in X'\}$.

(4) *if* (s,c) *blocks* μ' *and* $\mu'(c)$ *is non-empty, then* $[\arg\min_{P'_c} \mu'(c)] P'_c s$.

Proof. Part (1): For $c' \neq c$, c' only accepts students who are acceptable in each step of the SOSM and Algorithm 1. c rejects every student who is unacceptable under P'_c at the outset of Algorithm 1, and accepts no other student by the assumption that Algorithm 1 does not return to c for any subset of $\mu(c)$ as an input. Therefore μ' is individually rational.

Part (2): Suppose that for some $s \in S$ and $c' \in C$ such that $c' \neq c$, $c' P_s \mu'(s)$. Then, by the definition of the SOSM and Algorithm 1, s is rejected by c' either during the SOSM or in Algorithm 1. This implies that $|\mu'(c')| = q_{c'}$ and $[\arg\min_{P_{c'}} \mu'(c)] P_{c'} s$, implying that s and c' do not block μ'.

Part (3): Suppose $c P_s \mu'(s)$ for some $s \in S$. As in Part (2), this implies that s is rejected by c either during the SOSM or in Algorithm 1. Since Algorithm 1 does not return to c by assumption, s is rejected either during the SOSM or at the beginning of Algorithm 1. In the former case, if s is rejected during the SOSM, then $[\arg\min_{P_c} \mu(c)] P_c s$. In the latter case, (s,c) is not a blocking pair because s is declared unacceptable under P'_c.

Part (4): From Part (3), we have that

$$[\arg\min_{P_c} \mu(c)] P_c s. \tag{4}$$

By definition, $\mu'(c) \subseteq \mu(c)$ and expression (4) imply

$$[\arg\min_{P_c} \mu'(c)] P_c s, \tag{5}$$

when $\mu'(c)$ is non-empty.

When $\mu'(c)$ is non-empty, $\mu'(c)$ is constructed by including only acceptable students in $\mu(c)$ according to P'_c, so

$$[\arg\min_{P'_c} \mu'(c)] P'_c \emptyset. \tag{6}$$

Expressions (5) and (6) and the definition of a dropping strategy yield

$$[\arg\min_{P'_c} \mu'(c)] P'_c s.$$

□

Lemma 3 (Rejection chains)**.** *For any market and any* $c \in C$*, if Algorithm 1 does not return to* c *for any non-empty* $B^1_c \subseteq \mu(c)$*, then* c *cannot profitably manipulate by a dropping strategy.*

Proof. Consider an arbitrary dropping strategy (P'_c, q_c) and let $B^1_c = \{s \in \mu(c) | \emptyset P'_c s\}$. When Algorithm 1 does not return to c, we show that $\mu = \phi(P,q)$ is weakly preferred by college c to $\phi(P'_c, P_{-c}, q)$. Let $(P', q) = (P'_c, P_{-c}, q)$.

Let μ' be the matching resulting from Algorithm 1 with input B_c^1. In the first step of the proof, we construct a new matching μ'' by satisfying the blocking pairs involving college c such that college c weakly prefers μ over μ'' according to profile (P, q). The second step of the proof shows that μ'' is weakly preferred to $\phi(P', q)$ by college c. These two steps together will yield our desired conclusion.

First, suppose that c does not block μ' in (P', q). Then let $\mu'' = \mu'$. It is clear that college c weakly prefers μ over μ'' because $\mu''(c) = \mu'(c) \subseteq \mu(c)$. Otherwise, if c blocks μ', construct μ'' as follows: College c admits its most preferred students under P'_c who are willing to be matched possibly leaving the seats occupied by these students at other colleges vacant; that is,

$$\mu''(c') = \begin{cases} \mu'(c) \cup \arg\max_{(P'_c, q_c - |\mu'(c)|)} \{s \in S | c P_s \mu'(s)\} & c' = c, \\ \mu'(c') \setminus \mu''(c) & c' \neq c, \end{cases}$$

where $\arg\max_{(P'_c, q_c - |\mu'(c)|)} X$ denotes at most $q_c - |\mu'(c)|$ students that are most preferred under P'_c in set X. Recall that $\mu'(c) \subseteq \mu(c)$ by definition. Moreover, Part (3) of Lemma 2 shows that, under μ'', $|\mu'(c)|$ positions of c are filled with $\mu'(c)$ and the remaining $|\mu''(c) - \mu'(c)|$ positions are filled with students less preferred to $\arg\min_{P_c} \mu(c)$. Since preferences are responsive, we obtain

$$\mu(c) \succeq_c \mu''(c). \tag{7}$$

In the second step of the proof, we demonstrate that μ'' is weakly preferred to $\phi(P', q)$ by college c. The proof works by comparing μ'' to a stable matching μ''' in (P', q), that we construct below:

Case 1: If μ'' is stable in (P', q), then let $\mu''' = \mu''$. In this case, it is obvious that $\mu''(c)$ is weakly preferred to $\mu'''(c)$ under P'_c.

Case 2: Otherwise, observe the following properties of μ'':

1) Matching μ'' is individually rational.
2) College c is not part of a blocking pair under μ''.
3) The only blocking pairs in μ'' involve colleges who have vacant seats.

Property (1) follows by construction. To establish Property (2), we consider two cases. When $\mu'(c)$ is empty, $\mu''(c)$ is college c's most preferred students according to P'_c who are part of blocking pairs of μ'. College c is not part of a blocking pair under μ''. When $\mu'(c)$ is not empty, suppose college c is preferred by student s to $\mu''(s)$. In this case, Part (4) of Lemma 2 implies that s is less preferred than any student in $\mu'(c)$ under $P_{c'}$ and is less preferred than students in $\arg\max_{(P'_c, q_c - |\mu'(c)|)} \{s \in S | c P_s \mu'(s)\}$ by construction.

Therefore, college c does not form a blocking pair with this student, and, hence, c is not part of any blocking pair under μ''. Finally, Property (3) follows by construction and by Part (2) of Lemma 2.

With these properties in hand, we construct μ'''. Let $\mu''_0 \equiv \mu''$. μ''_0 may have a blocking pair involving a college with a vacant seat. Construct μ''_1 as follows. Let C_0 be the set of colleges that can be part of a blocking pair in μ''_0. Pick some $c' \in C_0$. Let student s be the student that college c' prefers the most within the set of students involved in blocking pairs with c'. Construct μ''_1 by assigning college c' to s, and if $\mu''_0(s) \neq \emptyset$, then s rejects $\mu''_0(s)$, leaving a vacant seat at college $\mu''_0(s)$. Under μ''_1, student s receives a strictly preferred assignment and each other student's assignment is unchanged. For college c, $\mu''_1(c)$ is weakly less preferred under P'_c than the initial matching, $\mu''_0(c)$, because the students are better off in μ''_1. Matching μ''_1 is individually rational, and as with μ''_0, the only blocking pair of μ''_1 involves colleges with vacant seats. If there are blocking pairs of μ''_1, then we repeat the same procedure to construct μ''_2, and so on.

At each repetition, the new matching is individually rational and each blocking pair of the new matching involves a college with a vacant seat. As a result, no additional students are rejected when the blocking pair is satisfied, and in each repetition, one student strictly improves and the remaining students do not change their assignment. Since students may improve their assignment a finite number of times, the procedure ends in finite time.

The ultimate matching μ''' is stable in (P', q) because it is individually rational and there are no blocking pairs. Moreover, for college c, each new matching is weakly less preferred under P'_c than the initial matching $\mu''_0(c)$ as students weakly improve in each new matching. Hence, $\mu''(c)$ is weakly preferred to $\mu'''(c)$ under P'_c.

Since the matching produced by SOSM is the least preferred stable matching of every college (attributed to Conway in Knuth (1976)), $\phi(P', q)(c)$ is weakly less preferred to the stable matching μ''' by college c under P'_c. This implies that $\phi(P', q)(c)$ is weakly less preferred to $\mu''(c)$ under P'_c, and since P'_c is a dropping strategy of P_c,

$$\mu''(c) \succeq_c \phi(P', q)(c). \tag{8}$$

Equations (7) and (8) together allow us to conclude that

$$\mu(c) \succeq_c \phi(P', q)(c),$$

showing that (P'_c, q_c) is not a profitable strategy when Algorithm 1 does not return to c. □

B.1.3 Lemma 7: Vanishing market power

We are interested in how often Algorithm 1 returns to a particular college c for the case where students draw their preferences from distribution $\mathcal{D}^n$. Let

$$\pi_c = \Pr[\text{Algorithm 1 returns to c for some } B_c^1 \subseteq \mu(c)].$$

Since Algorithm 1 returns to c for some B_c^1 whenever c can manipulate the SOSM (Lemmas 1 and 3), π_c gives an upper bound of the probability that c can manipulate the SOSM when others are truthful conditional on μ being realized as the matching under the SOSM. Here we will show Lemma 7, which bounds π_c for most colleges in large markets.

Consider the following algorithm, which is a stochastic variant of the SOSM.[44]

Algorithm 2. Stochastic Student-Optimal Gale-Shapley Algorithm

(1) Initialization: Let $l = 1$. For every $s \in S$, let $A_s = \emptyset$.

(2) Choosing the applicant:

  (a) If $l \leq |S|$, then let s be the lth student and increment l by one.[45]

  (b) If not, then terminate the algorithm.

(3) Choosing the applied:

  (a) If $|A_s| \geq k$, then return to Step 2.

  (b) If not, select c randomly from distribution $\mathcal{D}^n$ until $c \notin A_s$, and add c to A_s.

(4) Acceptance and/or rejection:

  (a) If c prefers each of its current mates to s and there is no vacant position, then c rejects s. Go back to Step 3.

  (b) If c has a vacant position or it prefers s to one of its current mates, then c accepts s. Now if c had no vacant position before accepting s, then c rejects the least preferred student among those who were matched to c. Let this student be s and go back to Step 3. If c had a vacant position, then go back to Step 2.

[44] To be more precise this is a stochastic version of the algorithm proposed by McVitie and Wilson (1970), which they show produces the same matching as the original SOSM proposed by Gale and Shapley (1962).

[45] Recall that students are ordered in an arbitrarily fixed manner.

A_s records colleges that s has already drawn from $\mathcal{D}^n$. When $|A_s| = k$ is reached, A_s is the set of colleges acceptable to s.

Under the SOSM, a student's application to her t^{th} most preferred college is independent of her preferences after $(t+1)^{\text{th}}$ choice on. Therefore the above algorithm terminates, producing the student-optimal stable matching of any realized preference profile which would follow from completing the draws for random preferences. Let μ be the student-optimal stable matching obtained by the above algorithm.

Suppose that Algorithm 2 is run and the stable matching μ is obtained. Now fix a college $c \in C$ and let B_c^1 be an arbitrary subset of $\mu(c)$. The **stochastic rejection chains** associated with B_c^1 is defined as follows. As the name suggests, this is a stochastic version of Algorithm 1.

Algorithm 3. Stochastic Rejection Chains

(1) Initialization:

  (a) Keep all the preference lists generated in Algorithm 2. Also, for each $s \in S$, let A_s be the set generated at the end of Algorithm 2. Let the student-optimal matching μ be the initial match of the algorithm. Let B_c^1 be a given subset of $\mu(c)$. Let $i = 0$. Let c reject all the students in B_c^1.

(2) Increment i by one.

  (a) If $B_c^i = \emptyset$, then terminate the algorithm.

  (b) If not, let s be the least preferred student by c among B_c^i, and let $B_c^{i+1} = B_c^i \setminus s$.

  (c) Iterate the following steps (call this iteration "Round i".)

    i. Choosing the applied:

      A. If $|A_s| \geq k$, then finish the iteration and go back to the beginning of Step 2.

      B. If not, select c' randomly from distribution $\mathcal{D}^n$ until $c' \notin A_s$, and add c' to A_s. If c is selected, terminate the algorithm.

    ii. Acceptance and/or rejection:

      A. If c' prefers each of its current mates to s and there is no vacant position, then c' rejects s; go back to the beginning of Step 2c.

      B. If c' has a vacant position or it prefers s to one of its current mates, then c' accepts s. Now if c' had no vacant position before accepting s, then c' rejects the least preferred student among those who were matched to c'. Let this rejected student be s and go back to the beginning of Step 2c. If c' had a vacant position, then finish the iteration and go back to the beginning of Step 2.

Algorithm 3 terminates either at Step 2a or at Step 2(c)iB. Similarly to Algorithm 1, we say that Algorithm 1 **returns to c** if it terminates at Step 2(c)iB and **does not return to c** if it terminates at Step 2a.

We are interested in how often the algorithm returns to c, as a student draws c from distribution $\mathcal{D}^n$. It is clear that the probability that Algorithm 1 returns to c is equal to the probability that Algorithm 3 returns to c. That is,

$$\pi_c = \Pr[\text{Algorithm 3 returns to } c \text{ for some } B_c^1 \subseteq \mu(c)].$$

This latter expression is useful since we can investigate the procedure step by step, utilizing conditional probabilities and conditional expectations. Recall that we have ordered the colleges in terms of decreasing popularity. Our notation is that when $c' \leq c$, we mean that c' is more popular than c, or $p_{c'} \geq p_c$ and when we write $c > m$, where m is a natural number, we mean the index of college c is larger than m. Let

$$V_c = \{c' \in C^n | c' \leq c, c' \notin A_s \text{ for every } s \in S^n \text{ at the end of Algorithm 2}\}, \text{ and}$$
$$Y_c = |V_c|.$$

V_c is a random set of colleges that are more popular than c ex ante but listed on no student's preference list at the end of Algorithm 2. Y_c is a random variable indicating the number of such colleges.[46]

Lemma 4. *For any $c > 4k$, we have*

$$E[Y_c] \geq \frac{c}{2} e^{-8\bar{q}nk/c}.$$

Proof. Let $Q^n = \sum_{c=1}^{k} p_c^n$. Then the probability that c' is not a student's ith choice, denoted $c_{(i)}$, given her first $(i-1)$ choices $c_{(1)}, \ldots, c_{(i-1)}$ is bounded as follows:

$$1 - \frac{p_{c'}^n}{1 - \sum_{j=1}^{i-1} p_{c_{(j)}}^n} \geq 1 - \frac{p_{c'}^n}{1 - Q^n}.$$

Let $E_{c'}$ be the event that $c' \notin A_s$ for every $s \in S$ at the end of Algorithm 2. Since there are at most $\bar{q}nk$ draws from $\mathcal{D}^n$ in Algorithm 2, the above inequality implies that

$$\Pr(E_{c'}) \geq \left(1 - \frac{p_{c'}^n}{1 - Q^n}\right)^{\bar{q}nk}.$$

[46] We abuse notation and denote a random variable and its realization by the same letter when there is no confusion.

If $c' > k$, there are at least $c' - k$ colleges that are at least as popular as c', but not among the k most popular colleges, so we obtain

$$p_{c'}^n \leq \frac{1 - Q^n}{c' - k}.$$

The last two inequalities imply

$$\Pr(E_{c'}) \geq \left(1 - \frac{1}{c' - k}\right)^{\bar{q}nk}. \tag{9}$$

We now show that for any $c' > 2k$,

$$\left(1 - \frac{1}{c' - k}\right)^{\bar{q}nk} \geq e^{-2\bar{q}nk/(c'-k)}. \tag{10}$$

To see this, first note that

$$\left(1 - \frac{1}{c' - k}\right)^{\bar{q}nk} \geq e^{-2\bar{q}nk/(c'-k)} \iff 1 - \frac{1}{c' - k} - e^{-2/(c'-k)} \geq 0.$$

Now, define a function $f(x) = 1 - x - e^{-2x}$. This function f is concave, and $f(0) = 0$ and $f(1/2) = 1/2 - 1/e > 0$. Therefore $f(x) \geq 0$ for any $x \in [0, 1/2]$. Since $c' > 2k$ and k is a positive integer, we have $c' - k > k \geq 1$. Since $c' - k$ is an integer, we thus obtain $c' - k \geq 2$ and hence $1/(c' - k) \in [0, 1/2]$. Therefore $1 - 1/(c' - k) - e^{-2/(c'-k)} = f(1/(c' - k)) \geq 0$, establishing inequality (10).

Moreover, for any $c' > 2k$,

$$e^{-2\bar{q}nk/(c'-k)} \geq e^{-4\bar{q}nk/c'}. \tag{11}$$

Combining inequalities (9), (10), and (11), we obtain:

$$\Pr(E_{c'}) \geq e^{-4\bar{q}nk/c'}.$$

Using the previous inequality, for any $c > 4k$, we have

$$E[Y_c] = \sum_{c'=1}^{c} \Pr(E_{c'}) \geq \sum_{c'=2k}^{c} e^{-4\bar{q}nk/c'} \geq \sum_{c'=c/2}^{c} e^{-8\bar{q}nk/c} = \frac{c}{2} e^{-8\bar{q}nk/c}.$$

□

For $B_c^1 \subseteq \mu(c)$, let

$$\pi_c^{B_c^1} = \Pr[\text{Algorithm 3 with input } B_c^1 \text{ returns to c}|Y_c > E[Y_c]/2, \mu].$$

$\pi_c^{B_c^1}$ gives an upper bound of the probability that c can manipulate the SOSM when others are truthful, conditional on two events: μ is the realized matching under the SOSM with truthful preferences and there are not too small a number of colleges ($Y_c > E[Y_c]/2$) that are more popular than c and appear nowhere on students' preference lists at the end of Algorithm 2.

Let $c^*(n) = 16\bar{q}nk/\ln(\bar{q}n)$. We will see, in Lemma 7, that $c^*(n)$ is the number of "very popular colleges" in a market with n colleges. Note that $c^*(n)/n$ converges to zero as $n \to \infty$, so the proportion of such colleges goes to zero. Except for these $c^*(n)$ colleges, the following lemma gives an upper bound for manipulability in a large market.

Lemma 5. *Suppose that n is sufficiently large and $c > c^*(n)$. Then we have*

$$\pi_c^{B_c^1} \leq \frac{4\bar{q}}{E[Y_c]},$$

for any $B_c^1 \subseteq \mu(c)$.

Proof. Consider Round 1, beginning with the least preferred student s of $B_c^1 \subseteq \mu(c)$ (if $B_c^1 = \emptyset$, then the inequality is obvious since $\pi_c^{B_c^1} = 0$.). Since $p_{c'}^n \geq p_c^n$ for any $c' \in V_c$, Round 1 ends at Step 2(c)iiB as a student applies to some college with vacant positions, at least with probability $1 - 1/(Y_c + 1) > 1 - 1/(E[Y_c]/2 + 1)$.

Now assume that all Rounds $1, \ldots, i-1$ end at Step 2(c)iiB. Then there are still at least $Y_c - (i-1)$ colleges more popular than c and with a vacant position, since at most $i-1$ colleges in V_c have had their positions filled at Rounds $1, \ldots, i-1$. Therefore Round i initiated by the least preferred student in B_c^i ends at Step 2(c)iiB with probability of at least $1 - 1/(E[Y_c]/2 - (i-1) + 1)$. Note that

$$E[Y_c]/2 - (i-1) + 1 \geq E[Y_c]/4 > 0, \tag{12}$$

for sufficiently large n. To see this, it is sufficient to show

$$E[Y_c]/4 - (i-1) + 1 \geq 0.$$

Since there are at most $\bar{q}$ rounds,

$$E[Y_c]/4 - (i-1) + 1 \geq E[Y_c]/4 - \bar{q} + 2.$$

The definition of Y_c implies that Y_c is weakly increasing in c. Lemma 4 provides a lower bound on $E[Y_c]$, which for $c > c^*(n)$ implies

$$E[Y_c]/4 - \bar{q} + 2 \geq \frac{2\bar{q}kn^{\frac{1}{2}}}{\ln(n)} - \bar{q} + 2,$$

which is positive for sufficiently large n, showing inequality (12) for sufficiently large n.

Since there are at most $\bar{q}$ rounds, Algorithm 3 fails to return to c with probability of at least

$$\prod_{i=1}^{\bar{q}} \left(1 - \frac{1}{E[Y_c]/2 - (i-1) + 1}\right) \geq \left(1 - \frac{1}{E[Y_c]/4}\right)^{\bar{q}}, \tag{13}$$

for sufficiently large n and $c > c^*(n)$, because of inequality (12).

Therefore we have that

$$\pi_c^{B_c^1} \leq 1 - \left(1 - \frac{1}{E[Y_c]/4}\right)^{\bar{q}} \leq \frac{4\bar{q}}{E[Y_c]},$$

where the last inequality holds since $1 - (1-x)^y \leq yx$ for any $x \in (0,1)$ and $y \geq 1$.[47] $\square$

We state without proof the following lemma (this is a straightforward generalization of Lemma 4.4 of Immorlica and Mahdian (2005)).

Lemma 6. *For every c, we have $Var[Y_c] \leq E[Y_c]$.*

Now we are ready to present and prove the last of the three key lemmas.

Lemma 7 (Vanishing market power)**.** *If n is sufficiently large and $c > c^*(n)$, then*

$$\pi_c \leq \frac{[\bar{q}(2^{\bar{q}} - 1) + 1]\ln(\bar{q}n)}{2k\sqrt{\bar{q}n}}.$$

Proof. By the fact that any probability is non-negative and less than or equal to one, the Chebychev inequality, and Lemma 6, we have

$$\begin{aligned} \Pr\left[Y_c \leq \frac{E[Y_c]}{2}\right] &\leq \Pr\left[Y_c \leq \frac{E[Y_c]}{2}\right] + \Pr\left[Y_c \geq \frac{3E[Y_c]}{2}\right] \\ &= \Pr\left[|Y_c - E[Y_c]| \geq \frac{E[Y_c]}{2}\right] \leq \frac{Var[Y_c]}{(E[Y_c]/2)^2} \leq \frac{4}{E[Y_c]}. \end{aligned}$$

[47] Note that conditions for this inequality is satisfied since $4/E[Y_c] \in (0,1)$ for any sufficiently large n and $c > c^*(n)$.

Since the probability of a union of events is at most the sum of the probabilities of individual events (Boole's inequality), Lemma 5 and the fact that there are at most $2^{\bar{q}} - 1$ non-empty subsets of $\mu(c)$ imply

$$\begin{aligned}
&\Pr[\text{Algorithm 3 returns to c for some } B_c^1 \subseteq \mu(c) | Y_c \geq E[Y_c]/2, \mu] \\
&\leq \sum_{B_c^1 \subseteq \mu(c)} \pi_c^{B_c^1} \\
&\leq \frac{4\bar{q}(2^{\bar{q}} - 1)}{E[Y_c]}.
\end{aligned}$$

This inequality holds for any matching μ. Therefore, we have the same upper bound for probability conditional on $Y_c > E[Y_c]/2$ but not on μ, that is,

$$\Pr\left[\text{Algorithm 3 returns to c for some } B_c^1 \subseteq \mu(c) | Y_c \geq \frac{E[Y_c]}{2}\right] \tag{14}$$

$$\leq \frac{4\bar{q}(2^{\bar{q}} - 1)}{E[Y_c]}. \tag{15}$$

By the above inequalities and the fact that probabilities are less than or equal to one,

$$\begin{aligned}
\pi_c &\leq \Pr\left[Y_c \leq \frac{E[Y_c]}{2}\right] + \Pr\left[Y_c > \frac{E[Y_c]}{2}\right] \times \frac{4\bar{q}(2^{\bar{q}} - 1)}{E[Y_c]} \\
&\leq \frac{4}{E[Y_c]} + \frac{4\bar{q}(2^{\bar{q}} - 1)}{E[Y_c]} \\
&= \frac{4[\bar{q}(2^{\bar{q}} - 1) + 1]}{E[Y_c]}.
\end{aligned}$$

Applying Lemma 4 and noting that $E[Y_c]$ is increasing in c so $E[Y_{c^*(n)}] \leq E[Y_c]$ for any $c > c^*(n) = 16\bar{q}nk/\ln(\bar{q}k)$, we complete the proof of Lemma 7.[48] □

B.1.4 Theorem 1

Now we prove Theorem 1. Let

$$\begin{aligned}
\alpha(n) = E[\#\{c \in C | \phi(S, C, (P'_c, P_{-c}), (q'_c, q_{-c})) \succ_c \phi(S, C, P, q) \\
\text{for some } (P'_c, q'_c) \text{ in the induced market}\} | \tilde{\Gamma}^n],
\end{aligned}$$

be the expected number of colleges that can manipulate in the market induced by random market $\tilde{\Gamma}^n$ under ϕ when others report preferences truthfully. By Lemma 1, it suffices to consider

[48]Note that Lemma 4 can be applied since for sufficiently large n and $c \geq c^*(n)$, we have $c > 4k$.

dropping strategies. By Lemma 3, the probability that $c \in C$ can successfully manipulate by some dropping strategy is at most π_c. Thus we obtain

$$\begin{aligned}
\alpha(n)/n &= \left[\sum_{c \in C^n} \Pr[c \text{ successfully manipulates}]\right]/n \\
&\leq c^*(n)/n + [\sum_{c \geq c^*(n)}^{n} \pi_c]/n \\
&\leq c^*(n)/n + [\sum_{c \geq c^*(n)}^{n} \frac{[\bar{q}(2^{\bar{q}}-1)+1]\ln(\bar{q}n)}{2k\sqrt{\bar{q}n}}]/n && \text{(by Lemma 7)} \\
&\leq \frac{16\bar{q}k}{\ln(\bar{q}n)} + \frac{(\bar{q}(2^{\bar{q}}-1)+1)\ln(\bar{q}n)}{2\sqrt{\bar{q}}k\sqrt{n}} && \text{(by } c^*(n) = 16\bar{q}nk/\ln(\bar{q}n)\text{)}.
\end{aligned}$$

The first term is proportional to $1/\ln(\bar{q}n)$ and the second term is proportional to $\ln(\bar{q}n)/\sqrt{n}$. Since both expressions approach zero as n approaches infinity, we obtain $\alpha(n)/n \to 0$ as $n \to \infty$, completing the proof.

B.2 Proof of Theorem 3

In this section, we consider the possibility of pre-arrangement, so we reintroduce S and C as arguments to the mechanism to avoid confusion.

B.2.1 Lemma 8

The following lemma says that a student that is involved in pre-arrangement is less preferred by the college to any student who is matched to it without pre-arrangement.

Lemma 8. *If $c \in C$ can manipulate via pre-arrangement with $s \in S$, then*

$$s' P_c s \text{ for every } s' \in \phi(S, C, P, q)(c).$$

Proof. Let $\mu(c) = \phi(S, C, P, q)(c)$. Theorem 2 of Sönmez (1999) implies that, for any stable mechanism, if c can manipulate via pre-arrangement with student s, then either $s \in \mu(c)$ or $s' P_c s$ for every $s' \in \mu(c)$. To show $s \notin \mu(c)$, suppose on the contrary that $s \in \mu(c)$. Consider matching μ' given by

$$\mu'(c') = \begin{cases} \mu(c) \setminus s & \text{if } c' = c, \\ \mu(c') & \text{otherwise.} \end{cases}$$

It is easy to see, from stability of μ in (S, C, P, q), that μ' is stable in $(S \setminus s, C, P_{-s}, q_c - 1, q_{-c})$.

Since the matching under the SOSM is weakly less preferred to any stable matching by colleges and preferences are responsive,

$$\begin{aligned}\phi(S, C, P, q)(c) &= \mu'(c) \cup s \\ &\succeq_c \phi(S \setminus s, C, P_{-s}, q_c - 1, q_{-c})(c) \cup s.\end{aligned}$$

Therefore c cannot manipulate the student-optimal stable mechanism via pre-arrangement. This is a contradiction, completing the proof. □

To profitably manipulate, a college has to pre-arrange a match with a strictly less preferred student. Then the disadvantage of being matched with a less desirable student should be compensated by matching to a better set of students in the centralized matching mechanism after pre-arrangement.

B.2.2 Theorem 3

Now we prove our result on pre-arrangement, Theorem 3. Start with matching $\phi(S, C, P, q)$, and consider what happens when college c reduces capacity by one. There are two cases to consider. First, the capacity reduction may not affect the matching of college c,

$$\phi(S, C, P, q_c - 1, q_{-c})(c) = \phi(S, C, P, q)(c).$$

This happens when $|\phi(S, C, P, q)(c)| < q_c$, or the number of students assigned to college c is less than its total capacity. Because there is an extra seat at college c, no student s would prefer to be assigned to college c over her matching $\phi(S, C, P, q)(s)$ because there would be a blocking pair, which contradicts the stability of $\phi(S, C, P, q)(s)$. Therefore, in the first case, pre-arrangement is not successful.

Second, the capacity reduction may affect the matching of college c. In this case, consider the rejection chain algorithm (Algorithm 1) starting with $\phi(S, C, P, q)$ with input $B_c^1 = \arg\min_{P_c} \phi(S, C, P, q)(c)$. We focus on the case where the rejection chain algorithm does not return to college c. Denote the resulting matching by μ'. Under μ', college c obtains $\phi(S, C, P, q)(c) \setminus \arg\min_{P_c} \phi(S, C, P, q)(c)$.

We claim that μ' is stable in $(S, C, P, q_c - 1, q_{-c})$. The ideas follow from Lemma 2.

First, we claim μ' is individually rational. For $c' \neq c$, c' only accepts students who are acceptable in each step of the SOSM and Algorithm 1. c is matched to $\phi(S, C, P, q)(c) \setminus \arg\min_{P_c} \phi(S, C, P, q)(c)$ at μ', which is clearly individually rational for c under $(P_c, q_c - 1)$. Therefore, μ' is individually rational.

Second, there is no blocking pair of μ' involving a college other than c. Suppose that for some $s \in S$ and $c' \in C$ such that $c' \neq c$, $c' P_s \mu'(s)$. Then, by the definition of the SOSM and Algorithm 1, s is rejected by c' either during the SOSM or in Algorithm 1. This implies that $|\mu'(c')| = q_{c'}$ and $[\arg\min_{P_{c'}} \mu'(c)]P_{c'}s$, implying that s and c' do not block μ'.

Third, there is no blocking pair of μ' involving college c. Suppose $cP_s\mu'(s)$ for some $s \in S$. As in the previous paragraph, this implies that s is rejected by c either during the SOSM or in Algorithm 1. Since Algorithm 1 does not return to c by assumption, s is rejected either during the SOSM or at the beginning of Algorithm 1. In the former case, the fact that s is rejected during the SOSM implies that $[\arg\min_{P_c} \mu(c)]P_c s$, so when college c has capacity $q_c - 1$, s will not form a blocking pair with student s because the other $q_c - 1$ students who are matched to c are more preferred to $[\arg\min_{P_c} \mu(c)]$. In the latter case, (s, c) is not a blocking pair because the student rejected at the beginning of Algorithm 1 is $\arg\min_{P_c} \phi(S, C, P, q)(c)$ by our assumption, the other $q_c - 1$ students who are matched to c are more preferred by college c, and the capacity of college c is $q_c - 1$.

Therefore, μ' is stable in $(S, C, P, q_c - 1, q_{-c})$.

The least preferred stable matching for colleges in $(S, C, P, q_c - 1, q_{-c})$ is equal to $\phi(S, C, P, q_c - 1, q_{-c})(c)$. Therefore, we can conclude that

$$\mu'(c) = \phi(S, C, P, q)(c) \setminus \arg\min_{P_c} \phi(S, C, P, q)(c) \succeq_c \phi(S, C, P, q_c - 1, q_{-c})(c). \tag{16}$$

Since every college is made weakly better off under the SOSM when the set of participating students increases (Gale and Sotomayor (1985)), we obtain

$$\phi(S, C, P, q_c - 1, q_{-c}) \succeq_c \phi(S \setminus s, C, P_{-s}, q_c - 1, q_{-c}). \tag{17}$$

Lemma 7 shows that the rejection chain algorithm (Algorithm 1) does not return with probability at least $1 - \pi_c$. As a result, with probability $1 - \pi_c$,

$$\begin{aligned}
\phi(S, C, P, q)(c) &\succeq_c \phi(S, C, P, q_c - 1, q_{-c})(c) \cup \arg\min_{P_c} \phi(S, C, P, q)(c) \\
&\succeq_c \phi(S \setminus s, C, P_{-s}, q_c - 1, q_{-c})(c) \cup \arg\min_{P_c} \phi(S, C, P, q)(c) \\
&\succeq_c \phi(S \setminus s, C, P_{-s}, q_c - 1, q_{-c})(c) \cup s,
\end{aligned}$$

where the first relation follows from (16) and responsiveness of preferences, the second relation follows from (17) and responsiveness, and the last relationship follows from Lemma 8 and responsiveness.

Therefore the probability that c benefits via pre-arrangement is at most π_c. Finally, by Lemma 7 we complete the proof (this last argument is similar to the one for Theorem 1 and hence omitted).

B.3 Proof of Theorems 2 and 5

Since Theorem 5 is a multi-region generalization of Theorem 2, we prove only the former.

B.3.1 Lemma 10: Uniform vanishing market power

We have a variant of Lemma 7 under the sufficient thickness assumption, which plays a crucial role in the proof of the theorems.

For $B_c^1 \subseteq \mu(c)$, let

$$\pi_c^{B_c^1} = \Pr[\text{Algorithm 3 associated with } B_c^1 \text{ returns to c}|Y_T(n) > E[Y_T(n)]/2, \mu].$$

First we show a variant of Lemma 5.

Lemma 9. *Suppose $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is regular and sufficiently thick. Let T be such that $E[Y_T(n)] \to \infty$ as $n \to \infty$. Suppose that n is sufficiently large. Then we have*

$$\pi_c^{B_c^1} \leq \frac{4T\bar{q}}{E[Y_T(n)]},$$

for any c and $B_c^1 \subseteq \mu(c)$.

Proof. Consider Round 1, beginning with the least preferred student s of $B_c^1 \subseteq \mu(c)$ (if $B_c^1 = \emptyset$, then the inequality is obvious since $\pi_c^{B_c^1} = 0$.). Since $p_{c'}^n(r) \geq p_c^n(r)/T$ for any $c' \in V_T(n)$ and $r = 1, \dots, R$, Round 1 ends at 2(c)iiB as a student applies to some college with vacant positions, at least with probability $1 - 1/(Y_T(n)/T + 1) > 1 - 1/(E[Y_T(n)]/2T + 1)$.

Now assume that all Rounds $1, \dots, i-1$ end at Step 2(c)iiB. Then there are still at least $Y_T(n) - (i-1)$ colleges more popular than c and with a vacant position, since at most $i-1$ colleges in $V_T(n)$ have had their positions filled at Rounds $1, \dots, i-1$. Therefore Round i initiated by the least preferred student in B_c^i ends at Step 2(c)iiB with probability of at least $1 - 1/(E[Y_T(n)]/2T - (i-1) + 1)$. Since there are at most $\bar{q}$ rounds, Algorithm 3 fails to return to c with probability of at least

$$\begin{aligned}\prod_{i=1}^{\bar{q}} \left(1 - \frac{1}{E[Y_T(n)]/2T - (i-1) + 1}\right) &\geq \left(1 - \frac{1}{E[Y_T(n)]/2T - \bar{q} + 2}\right)^{\bar{q}} \\ &\geq \left(1 - \frac{1}{E[Y_T(n)]/4T}\right)^{\bar{q}}.\end{aligned}$$

The first inequality follows since $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is sufficiently thick, n is sufficiently large and $i \leq \bar{q}$ for each i. The second inequality holds since $E[Y_T(n)]/2 - \bar{q} \geq E[Y_T(n)]/4 > 0$, which follows since $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is sufficiently thick and n is sufficiently large. Therefore we have that

$$\pi_c^{B_c^1} \leq 1 - \left(1 - \frac{1}{E[Y_T(n)]/4T}\right)^{\bar{q}}$$
$$\leq \frac{4T\bar{q}}{E[Y_T(n)]},$$

where the last inequality holds since $1-(1-x)^y \leq yx$ for any $x \in (0,1)$ and $y \geq 1$. □

Lemma 10 (Uniform vanishing market power)**.** *Suppose that* $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ *is regular and sufficiently thick. For any sufficiently large* n *and any* $c \in C$*, we have*

$$\pi_c \leq \frac{4\left[T\bar{q}(2^{\bar{q}}-1)+1\right]}{E[Y_T(n)]}.$$

Proof. By Lemma 9 and an argument similar that which leads to expression (14) in Lemma 7, we obtain

$$\Pr[\text{Algorithm 3 returns to c}|Y_T(n) > E[Y_T(n)]/2] \leq \frac{4T\bar{q}(2^{\bar{q}}-1)}{E[Y_T(n)]}.$$

Therefore we have

$$\pi_c \leq \Pr\left[Y_T(n) \leq E[Y_T(n)]/2\right] + \Pr\left[Y_T(n) > E[Y_T(n)]/2\right] \times \frac{4T\bar{q}(2^{\bar{q}}-1)}{E[Y_T(n)]}$$
$$\leq \frac{4}{E[Y_T(n)]} + \frac{4T\bar{q}(2^{\bar{q}}-1)}{E[Y_T(n)]}$$
$$\leq \frac{4\left[T\bar{q}(2^{\bar{q}}-1)+1\right]}{E[Y_T(n)]},$$

completing the proof. □

B.3.2 Theorems 2 and 5

We only prove Theorem 5, since Theorem 2 is a special case when $R = 1$. Suppose that colleges other than c are truth-telling, that is, any $c' \neq c$ reports $(P_{c'}, q_{c'})$. Lemmas 1 and 3 apply here since they do not rely on assumptions about student preferences. These lemmas imply that the probability that c profitably manipulates is at most π_c. By Lemma 10 and sufficient thickness, for any $\varepsilon > 0$, there exists n_0 such that for any market $\tilde{\Gamma}^n$ with $n > n_0$, we have

$$\Pr\left[u(\phi(S,C,P'_c,P_{-c},q)(c)) > u(\phi(S,C,P,q)(c)) \text{ for some } P'_c\right] < \frac{\varepsilon}{\bar{q}\sup_{n\in\mathbb{N}, s\in S^n, c\in C^n} u_c(s)}.$$

Such n_0 can be chosen independent of $c \in C^n$. For any $n > n_0$, for any $c \in C^n$ we have

$$\begin{aligned}
&Eu_c(\phi(S, C, (P'_c, P_{-c}), (q'_c, q_{-c}))(c)) - Eu_c(\phi(S, C, P, q)) \\
<&\Pr\left[u_c(\phi(S, C, (P'_c, P_{-c}), (q'_c, q_{-c}))(c) > u_c(\phi(S, C, P, q)(c))\right] \bar{q} \sup_{n \in \mathbb{N}, s \in S^n, c \in C^n} u_c(s) \\
<&\varepsilon,
\end{aligned}$$

which implies that truthful reporting is an ε-Nash equilibrium.

B.4 Proofs of Propositions 1 and 2

B.4.1 Proposition 1

Let a and T satisfy the condition of nonvanishing proportion of popular colleges. Let $c = [an]$. Then it is obvious that $V_c \subseteq V_T(n)$ and hence $Y_c \leq Y_T(n)$. For sufficiently large n, Lemma 4 shows that

$$E[Y_T(n)] \geq E[Y_c] \geq \frac{c}{2} e^{-8\bar{q}nk/c}.$$

$c = [an]$ implies that $\frac{c}{2} e^{-8\bar{q}nk/c} \to \infty$ as $n \to \infty$ with the order $O(n)$. Therefore $E[Y_T(n)] \to \infty$ as $n \to \infty$, completing the proof.

For Example 5, let $c = [\gamma n / \ln(n)]$. Substituting into $\frac{c}{2} e^{-8\bar{q}nk/c}$, we obtain

$$E[Y_T(n)] \geq \frac{c}{2} e^{-8\bar{q}nk/c} \geq \left[\frac{\gamma n / \ln(n) - 1}{2}\right] n^{\frac{-8\bar{q}k}{\gamma}} = \frac{\gamma n^{1 - \frac{8\bar{q}k}{\gamma}}}{2 \ln(n)} - \frac{1}{2} n^{-\frac{8\bar{q}k}{\gamma}}.$$

Therefore, for $\gamma > 8\bar{q}k$, then $n^{1 - \frac{8\bar{q}k}{\gamma}} / \ln(n) \to \infty$ and $n^{-\frac{8\bar{q}k}{\gamma}} \to 0$ as $n \to \infty$. Thus, we conclude that $E[Y_T(n)] \to \infty$ as $n \to \infty$.

B.4.2 Proposition 2

Let

$$V_T^*(n) = \{c \in C^n | p_1^n(r)/p_c^n(r) \leq T \quad \text{for all } r \in \{1, ..., \bar{R}\}\}.$$

Then we have $V_T(n) = \{c \in V_T^*(n) : |\{s \in S^n | c P_s s\}| < q_c\}$. Let $\eta_r(c) = |\{c' \in C^n | p_c^n(r) \leq p_{c'}^n(r)\}|$ be the order of c with respect to popularity in distribution $D^n(r)$. For example, if college c is the most popular among students in region 1 and the least popular among those in region 2, then $\eta_1(c) = 1$ and $\eta_2(c) = n$.

Part (1): Example 6. Let $T = 4$, for example. Then, $V_4^*(n) = \{n/4, n/4 + 1, \ldots, 3n/4\}$. Consider any college $c \in V_4^*(n)$. Let s belong to region $r \in \{1, 2\}$. Since s picks colleges k times according to $D^n(r)$, the probability that c does not appear in the preference list of student s, denoted by $\Pr(F_{c,s})$, is bounded as follows:

$$\Pr(F_{c,s}) \geq \left(1 - \frac{p_c^n(r)}{1 - Q^n(r)}\right)^k,$$

where

$$Q^n(r) = \sum_{c:\eta_r(c) \leq k} p_c^n(r).$$

For any sufficiently large n, we have that $\eta_r(c) > 2k$ for any $c \in V_4^*(n)$ and $r = 1, 2$ since $[n - 2k] > 3n/4 > c > n/4 > 2k$. For such colleges,

$$p_c^n(r) \leq \frac{1 - Q^n(r)}{\eta_r(c) - k} \leq \frac{1 - Q^n(r)}{\eta_r(c)/2}.$$

So

$$\Pr(F_{c,s}) \geq \left(1 - \frac{2}{\eta_r(c)}\right)^k.$$

Since $\eta_r(c) \geq n/4$ for any $c \in V_4^*(n)$ and any $r = 1, 2$, we have

$$\Pr(F_{c,s}) \geq \left(1 - \frac{8}{n}\right)^k.$$

Let E_c be the event that c is not listed by any student. Then, since students draw colleges independently, we have

$$\Pr(E_c) = \prod_{s \in S^n} \Pr(F_{c,s}) \geq (1 - 8/n)^{k\bar{q}n} \to e^{-8k\bar{q}},$$

as $n \to \infty$. Therefore,

$$E[Y_T(n)] = \sum_{c \in V_4^*(n)} \Pr(E_c) \geq \frac{n}{2}(1 - 8/n)^{k\bar{q}n} \to \infty,$$

as $n \to \infty$ (with the order $O(n)$), completing the proof.

Part (2): Example 7. As discussed in Example 7, for any colleges c and c' and region r, we have that

$$p_c^n(r)/p_{c'}^n(r) = \tilde{p}_{r(c)}(r)/\tilde{p}_{r(c')}(r) > 0.$$

Since there are only finite regions, $V_T^*(n) = C^n$ for any sufficiently large T. Fix such T.

As in the proof of Part (1), for any c and s we have

$$\Pr(F_{c,s}) \geq \left[1 - \frac{p_c^n(r(s))}{1 - Q^n(r(s))}\right]^k.$$

Since we have that $p_c^n(r)/p_{c'}^n(r) < T$ for any $c, c' \in C^n$,

$$\begin{aligned}\frac{p_c^n(r)}{1 - Q^n(r)} &\leq \frac{p_c^n(r)}{(n-k)p_c^n(r)/T} \\ &\leq \frac{2T}{n},\end{aligned}$$

for any sufficiently large n. So we have

$$\Pr(E_c) = \prod_{s \in S^n} \Pr(F_{c,s}) \geq \prod_{s \in S^n} \left(1 - \frac{2T}{n}\right)^k \geq (1 - 2T/n)^{k\bar{q}n} \to e^{-2k\bar{q}T},$$

as $n \to \infty$. Therefore

$$E[Y_T(n)] = \sum_{c \in C^n} \Pr(E_c)$$

approaches infinity with the order $O(n)$, completing the proof.

Remark 1. In Examples 4, 6 and 7, the order of convergence of $E[Y_T(n)]$ is $O(n)$. This implies that, by Lemma 10, the order of convergence of the probability of profitable manipulation is $O(1/n)$. This is the same order as in the uniform distribution case, analyzed by Roth and Peranson (1999) and Immorlica and Mahdian (2005).

B.5 Proof of Theorem 7: Coalitional manipulation

The proof is based on a series of arguments similar to those for Theorem 1. First, dropping strategies are exhaustive for manipulations involving a coalition of agents.

Lemma 11 (Dropping strategies are exhaustive for coalitional manipulations)**.** *Consider an arbitrary stable mechanism φ. Fix preference profile (P, q) and let $\bar{C} \subseteq C$ be a coalition of colleges. For some arbitrary report $(\tilde{P}_{\bar{C}}, \tilde{q}_{\bar{C}})$ by this coalition, suppose that $\varphi((\tilde{P}_{\bar{C}}, P_{-\bar{C}}), (\tilde{q}_{\bar{C}}, q_{-\bar{C}})) = \mu$. Then there exists a dropping strategy $(P'_{\bar{C}}, q_{\bar{C}}) = (P'_c, q_c)_{c \in \bar{C}}$ such that $\varphi((P'_{\bar{C}}, P_{-\bar{C}}), q) \succeq_c \mu$ for each $c \in \bar{C}$.*

Proof. For each $c \in \bar{C}$, construct dropping strategy (P'_c, q_c) such that P'_c lists all the students it is matched to under μ who are acceptable ($\{s \in \mu(c) | s P_c \emptyset\}$) in the same relative order as in P_c, and reports every other student as unacceptable. Let $(P', q) = (P'_{\bar{C}}, P_{-\bar{C}}, q)$.

We will show $\varphi(P', q)(c)$ is equal to $\{s \in \mu(c) | sP_c\emptyset\}$ for each $c \in \bar{C}$. This equality implies that $\varphi(P', q) \succeq_c \mu$ for each $c \in \bar{C}$, since $\varphi(P', q)(c)$ can only differ from $\mu(c)$ in having no unacceptable students under the true preference list P_c. The proof proceeds in two steps.

First, consider the matching μ' obtained from μ by having each $c \in \bar{C}$ keep only the students in $\mu(c)$ that are acceptable under P_c. That is,

$$\mu'(c) = \begin{cases} \{s \in \mu(c) | sP_c\emptyset\} & c \in \bar{C}, \\ \mu(c) & c \notin \bar{C}. \end{cases}$$

Consider the properties of μ' under (P', q). First, μ' is individually rational under (P', q). Second, there is no blocking pair involving any $c \in \bar{C}$, since $\mu'(c)$ is exactly the set of students acceptable under P'_c. However, there may be a blocking pair involving another college and students who are unmatched. In this case, we can define a procedure which ultimately yields a matching that is stable under (P', q):

Starting from $\mu'_0 \equiv \mu'$, let S_0 be the set of students that can be part of a blocking pair in μ'_0. Pick some $s \in S_0$. Let college c' be the college that the student s prefers the most within the set of colleges involved in blocking pairs with s. Construct μ'_1 by assigning student s to c' and if $|\mu'_0(c')| = q_{c'}$, then let c' reject its least preferred student. By construction, μ'_1 is individually rational, college c is matched to the same set of students, college c' strictly prefers $\mu'_1(c')$ over $\mu'_0(c')$ and the assignments of all other colleges are unchanged. As with μ'_0, the only blocking pair of μ'_1 involves unmatched students. If there are blocking pairs of μ'_1, then we repeat the same procedure to construct μ'_2, and so on.

At each repetition, the new matching is individually rational, each college $c \in \bar{C}$ is matched to the same set of students, one other college strictly improves and the remaining colleges are not worse off since each blocking pair involves an unassigned student. Since colleges can strictly improve only a finite number of times, this procedure terminates in finite time. The ultimate matching μ'' is stable in (P', q) because it is individually rational and there are no blocking pairs. Moreover, since each college $c \in \bar{C}$ obtains the same matching in each repetition,

$$\mu''(c) = \mu'(c), \text{ for each } c \in \bar{C}. \tag{18}$$

The second step of the proof utilizes the fact that for any college, the same number of students are matched to it across different stable matchings. Since μ'' is stable in (P', q) and $\varphi(P', q)$ is stable in (P', q) by definition, $|\mu''(c)| = |\varphi(P', q)(c)|$, for each $c \in \bar{C}$. Since for each $c \in \bar{C}$, there are just $|\mu''(c)|$ acceptable students under P'_c, this implies that

$$\mu''(c) = \varphi(P', q)(c), \text{ for each } c \in \bar{C}. \tag{19}$$

Equations (18) and (19) together imply that for each $c \in \bar{C}$,

$$\varphi(P', q)(c) = \mu'(c) = \{s \in \mu(c) | sP_c\emptyset\}.$$

As a result, for each $c \in \bar{C}$

$$\varphi(P', q)(c) \succeq_c \mu(c),$$

since $\varphi(P', q)(c)$ can only differ from $\mu(c)$ in having no unacceptable students under the true preference list P_c. □

B.5.1 Lemma 13: Rejection chains with coalitions

For preference profile (P, q), let μ be the student-optimal stable matching. Let B_c^1 be an arbitrary subset of $\mu(c)$. The **rejection chains** algorithm with input $(B_c^1)_{c \in \bar{C}}$ is defined as follows.

Algorithm 4. REJECTION CHAINS WITH COALITIONS

(1) Initialization: μ is the student-optimal stable matching μ, input $(B_c^1)_{c \in \bar{C}}$, $\bar{C}^1 = \bar{C}$, and let $j = 0$.

(2) Algorithm:

  (a) Increment j by one (iterate through colleges).

    i. If $\bar{C}^j = \emptyset$, then terminate the algorithm.

    ii. If not, pick some $c \in \bar{C}^j$, and let $\bar{C}^{j+1} = \bar{C}^j \setminus c$ and let $i = 0$.

  (b) Increment i by one (iterate through students)

    i. If $B_c^i = \emptyset$, then go to beginning of Step 2a.

    ii. If not, let s be the least preferred student by c among B_c^i, and let $B_c^{i+1} = B_c^i \setminus s$.

    iii. Iterate the following steps (call this iteration "Round i").

    Choosing the applied:

    A. If s has already applied to every acceptable college, then finish the iteration and go back to the beginning of Step 2b.

    B. If not, let c' be the most preferred college of s among those which s has not yet applied while running the SOSM or previously within this algorithm. If $c' \in \bar{C}$, terminate the algorithm.

    Acceptance and/or rejection:

    A. If c' prefers each of its current mates to s and there is no vacant position, then c' rejects s; go back to the beginning of Step 2(b)iii.

B. If c' has a vacant position or it prefers s to one of its current mates, then c' accepts s. Now if c' had no vacant position before accepting s, then c' rejects the least preferred student among those who were matched to c'. Let this rejected student be s and go back to the beginning of Step 2(b)iii. If c' had a vacant position, then finish the iteration and go back to the beginning of Step 2b.

Algorithm 4 terminates either at Step 2(a)i or at Step 2(b)iiiB. We say that Algorithm 4 **does not return to** $\bar{C}$ if it terminates at Step 2(a)i and it **returns to** $\bar{C}$ if it terminates at Step 2(b)iiiB.

Lemma 12. *Consider coalition $\bar{C}$ and suppose that under (P, q), Algorithm 4 with each possible collection of sets consisting of subsets of $\mu(c)$ for each $c \in \bar{C}$ as an input fails to return to $\bar{C}$. For any dropping strategy $(P'_{\bar{C}}, q_{\bar{C}})$, let $B^1_c = \{s \in \mu(c) | \emptyset P'_c s\}$ for each $c \in \bar{C}$ such $B^1_c \neq \emptyset$ for at least one $c \in \bar{C}$, and let μ' be a matching obtained at the end of Algorithm 4 with input $(B^1_c)_{c \in \bar{C}}$. Then, under $(P'_{\bar{C}}, P_{-\bar{C}}, q)$,*

(1) *μ' is individually rational,*

(2) *no $c \notin \bar{C}$ is a part of a blocking pair of μ', and*

(3) *for $c \in \bar{C}$, if (s, c) blocks μ', then $[\arg\min_{P_c} \mu(c)] P_c s$, and*

(4) *for $c \in \bar{C}$, if (s, c) blocks μ' and $\mu'(c)$ is non-empty, then $[\arg\min_{P'_c} \mu'(c)] P'_c s$.*

Proof. Part (1): For $c \notin \bar{C}$, c only accepts students who are acceptable in each step of the SOSM and Algorithm 4. Each $c \in \bar{C}$ rejects every student who is unacceptable under $P'_{\bar{C}}$ at the outset of Algorithm 4, and accepts no other student by the assumption that Algorithm 4 does not return to $c \in \bar{C}$ for any sequence of sets consisting of subsets of $\mu(c)$ for each $c \in \bar{C}$ as input. Therefore μ' is individually rational.

Part (2): Suppose that for some $s \in S$ and $c \in C \setminus \bar{C}$, $c P_s \mu'(s)$. Then, by the definition of the SOSM and Algorithm 4, s is rejected by c either during the SOSM or in Algorithm 4. This implies that $|\mu'(c)| = q_c$ and $[\arg\min_{P_c} \mu'(c)] P_c s$, implying that s and c do not block μ'.

Part (3): Suppose $c P_s \mu'(s)$ for some $s \in S$ and $c \in \bar{C}$. As in Part (2), this implies that s is rejected by c either during the SOSM or in Algorithm 4. Since Algorithm 4 does not return to c by assumption, s is rejected either during the SOSM or at the beginning of Algorithm 4. In the former case, if s is rejected during the SOSM, then $[\arg\min_{P_c} \mu(c)] P_c s$. In the latter case, (s, c) is not a blocking pair because s is declared unacceptable under P'_c.

Part (4): From Part (3), for each $c \in \bar{C}$, we have that

$$[\arg\min_{P_c} \mu(c)] P_c s. \tag{20}$$

By definition, $\mu'(c) \subseteq \mu(c)$ and expression (20) imply

$$[\arg\min_{P_c} \mu'(c)]P_c s, \tag{21}$$

when $\mu'(c)$ is non-empty.

When $\mu'(c)$ is non-empty, $\mu'(c)$ is constructed by including only acceptable students in $\mu(c)$ according to P'_c, so

$$[\arg\min_{P'_c} \mu'(c)]P'_c\emptyset. \tag{22}$$

Expressions (21) and (22) and the definition of a dropping strategy yield

$$[\arg\min_{P'_c} \mu'(c)]P'_c s.$$

□

Lemma 13 (Rejection chains with coalitions)**.** *For any market and any $\bar{C} \subseteq C$, if Algorithm 4 does not return to $\bar{C}$ with each possible collection of sets $(B^1_c)_{c\in\bar{C}}$, where B^1_c is a subset of $\mu(c)$ for each $c \in \bar{C}$, as input (with at least one B^1_c non-empty), then $\bar{C}$ cannot profitably manipulate by a dropping strategy.*

Proof. Consider an arbitrary dropping strategy $(P'_{\bar{C}}, q_{\bar{C}})$ and let $B^1_c = \{s \in \mu(c)|\emptyset P'_c s\}$ for each $c \in \bar{C}$. When Algorithm 4 does not return to $\bar{C}$, we show that $\mu = \phi(P, q)$ is weakly preferred by each college $c \in \bar{C}$ to $\phi(P'_{\bar{C}}, P_{-\bar{C}}, q)$. Let $(P', q) = (P'_{\bar{C}}, P_{-\bar{C}}, q)$.

Let μ' be the matching resulting from Algorithm 4 with input $(B^1_c)_{c\in\bar{C}}$. The first step of the proof constructs a new matching μ'' by satisfying the blocking pairs in μ' such that μ'' is stable in (P', q)).

We construct μ'' from μ' as follows. Parts (1) and (2) of Lemma 12 imply that μ' is individually rational and no $c \notin \bar{C}$ is part of a blocking pair of μ'. Parts (3) and (4) of Lemma 12 states that if there is a blocking pair involving some $c \in \bar{C}$, then the student involved in the blocking pair is less preferred than any student in $\mu(c)$ under P_c and less preferred than any student in $\mu'(c)$ under P'_c.

Let $\mu''_0 = \mu'$. Begin by selecting some college $c \in \bar{C}$ involved in a blocking pair of μ''_0. Construct matching μ''_1 by assigning college c its most preferred students in blocking pairs up to capacity. The resulting matching is individually rational, at least one student strictly improves over μ''_0, no other student is made worse off, and every blocking pair of μ''_1 involves a college with a vacant seat. Finally, since students are made weakly better off in μ''_1 than in μ''_0, it must be the case that if (s, c) blocks μ''_1, then $\arg\min_{P_c} \mu(c)P_c s$ and $\arg\min_{P'_c} \mu'(c)P'_c s$.

Next, we construct μ''_2 by selecting college c who is involved in a blocking pair of μ''_1. Note that this can either be a college in $\bar{C}$ or not, but the college that is selected must have a vacant seat. As before, we assign c the most preferred students it forms a blocking pair with up to capacity. As before, the resulting matching is individually rational, at least one student strictly improves over μ''_1, no other student is made worse off, and every blocking pair of μ''_2 involves a college with a vacant seat. Finally, since students are made weakly better in μ''_2 than in μ''_1, it must be the case that if (s, c) blocks μ''_2, then $\arg\min_{P_c} \mu(c) P_c s$ and $\arg\min_{P'_c} \mu'(c) P_c s$.

We repeat this process to construct μ''_3, and so on.

In each step of this construction, no additional students are rejected when we satisfy a blocking pair, and in each step, at least one student strictly improves and the remaining students are not made worse off. Since students may improve their assignment only a finite number of times, the procedure ends in finite time.

The ultimate matching μ'' is stable in (P', q) because it is individually rational and there are no blocking pairs. Moreover, for college c, each new student matched to college c is weakly less preferred under P'_c than any student in the initial matching $\mu''_0(c)$ as students obtain a weakly more preferred college in each step. Hence, $\mu'(c)$ is weakly less preferred to $\mu''(c)$ under P'_c. Also, for college c, each new student matched to college c is weakly less preferred under P_c than any student in the matching $\mu(c)$ because students obtain a weakly more preferred college in each step. Thus, $\mu''(c)$ is weakly less preferred to $\mu(c)$ under P_c.

Since the matching produced by SOSM is the least preferred stable matching of every college, $\phi(P', q)(c)$ is weakly less preferred to the stable matching μ'' by each college $c \in \bar{C}$ in (P', q). This will imply that $\phi(P', q)(c)$ is weakly less preferred to $\mu''(c)$ under P'_c for each $c \in \bar{C}$, and since P'_c is a dropping strategy of P_c,

$$\mu''(c) \succeq_c \phi(P', q)(c), \text{ for each } c \in \bar{C}.$$

□

B.5.2 Lemma 14: Uniform vanishing market power for coalitions

For the coalitions, we define

$$\pi_c = \Pr[\text{Algorithm 4 returns to } \bar{C} \text{ for some } (B^1_c)_{c \in \bar{C}} \subseteq \cup_{c \in \bar{C}} \mu(c)].$$

Lemma 14. *Suppose that $(\tilde{\Gamma}^1, \tilde{\Gamma}^2, \dots)$ is regular and sufficiently thick. For any sufficiently large n and coalition $\bar{C} \subseteq C$, we have*

$$\pi_{\bar{C}} \leq \frac{4\left[T\bar{q} \cdot |\bar{C}| \cdot (2^{\bar{q} \cdot |\bar{C}|} - 1) + 1\right]}{E[Y_T(n)]}.$$

Proof. The proof follows exactly the steps leading to Lemma 10 with two modifications. The first modification replaces the first instance of $\bar{q}$ in the expression in Lemma 10 with $\bar{q} \cdot |\bar{C}|$

because in the proof that corresponds to Lemma 9 we must allow for the possibility of $\bar{q}$ rounds for each of the $|\bar{C}|$ colleges. The second modification replaces $2^{\bar{q}}$ in the expression in Lemma 10 with $2^{\bar{q}\cdot|\bar{C}|}$ because in the proof that corresponds to Lemma 7, there are at most $2^{\bar{q}\cdot|\bar{C}|}-1$ non-empty subsets of $\cup_{c\in\bar{C}}\mu(c)$. □

Finally, the proof of Theorem 7 follows because in sufficiently thick markets $E[Y_T(n)]\to\infty$ as $n\to\infty$ and $|\bar{C}|\le m$, so $\pi_{\bar{C}}\to 0$.

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School Choice: A Mechanism Design Approach," *American Economic Review*, 2003, *93*, 729–747.

– , Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez, "The Boston Public School Match," *American Economic Review, Papers and Proceedings*, 2005, *95*, 368–371.

– , – , – , and – , "Changing the Boston School Choice Mechanism," 2006. NBER Working paper, 11965.

– , – , and – , "The New York City High School Match," *American Economic Review, Papers and Proceedings*, 2005, *95*, 364–367.

– , – , and – , "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match," 2008. forthcoming, *American Economic Review*.

Alcalde, José and Salvador Barberà, "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, 1994, *4*, 417–435.

Blum, Yosef, Alvin E. Roth, and Uriel Rothblum, "Vacancy Chains and Equilibration in Senior-Level Labor Markets," *Journal of Economic Theory*, 1997, *76*, 362–411.

Bogomolnaia, Anna and Hervé Moulin, "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 2001, *100*, 295–328.

Bulow, Jeremy and Jonathan Levin, "Matching and Price Competition," *American Economic Review*, 2006, *96*, 652–668.

Cantala, David, "Restabilizing Markets at Senior Level," *Games and Economic Behavior*, 2004, *48*, 1–17.

Cripps, Martin and Jeroen Swinkels, "Depth and Efficiency of Large Double Auctions," *Econometrica*, 2006, *74*, 47–92.

Day, Robert and Paul Milgrom, "Core-Selecting Auctions," *International Journal of Game Theory*, 2008, *36*, 393-407.

Demange, Gabrielle, David Gale, and Marilda A. O. Sotomayor, "A Further Note on the Stable Matching Problem," *Discrete Applied Mathematics*, 1987, *16*, 217–222.

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

Ehlers, Lars, "In Search of Advice for Participants in Matching Markets which use the Deferred Acceptance Algorithm," *Games and Economic Behavior*, 2004, *48*, 249–270.

Erdil, Aytek and Haluk Ergin, "What's the Matter with Tie-Breaking? Improving Efficiency in School Choice," *American Economic Review*, 2008, 98(3), 669-689.

Ergin, Haluk and Tayfun Sönmez, "Games of School Choice under the Boston Mechanism," *Journal of Public Economics*, 2006, *90*, 215–237.

Fudenberg, Drew, Markus Mobius, and Adam Szeidl, "Existence of Equilibria in

Large Double Auctions," *Journal of Economic Theory*, 2007, *133*, 550–567.
Gale, David and Lloyd S. Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69*, 9–15.
– **and Marilda A. Oliveira Sotomayor**, "Some Remarks on the Stable Matching Problem," *Discrete Applied Mathematics*, 1985, *11*, 223–232.
Gresik, Thomas and Mark Satterthwaite, "The Rate at Which a Simple Market Converges to Efficiency as the Number of Traders Increases," *Journal of Economic Theory*, 1989, *48*, 304–332.
Haeringer, Guillaume and Flip Klijn, "Constrained School Choice," 2007. UAB, Unpublished mimeo.
Hatfield, John and Paul Milgrom, "Matching with Contracts," *American Economic Review*, 2005, *95*, 913–935.
Immorlica, Nicole and Mohammad Mahdian, "Marriage, Honesty, and Stability," *SODA*, 2005, 53–62.
Kelso, Alexander and Vincent Crawford, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 1982, *50*, 1483–1504.
Kesten, Onur, "On Two Kinds of Manipulation for School Choice Problems," 2006. Unpublished mimeo.
Klaus, Bettina and Flip Klijn, "Stable Matchings and Preferences of Couples," *Journal of Economic Theory*, 2005, *121*, 75–106.
Knuth, Donald E., *Marriage Stables*, Montreal: Les Presse de l'Université de Montreal, 1976.
– **, Rajeev Motwani, and Boris Pittel**, "Stable Husbands," *Random Structures and Algorithms*, 1990, *1*, 1–14.
Kojima, Fuhito, "Mixed Strategies in Games of Capacity Manipulation in Hospital-Intern Markets," *Social Choice and Welfare*, 2006, *27*, 25–28.
– , "Finding All Stable Matchings with Couples," 2007a. mimeo.
– , "When can Manipulations be Avoided in Two-Sided Matching Markets? Maximal Domain Results," *Contributions to Theoretical Economics*, 2007b, *7*, Article 32.
– **and Mihai Manea**, "Strategy-Proofness of the Probabilistic Serial Mechanism in Large Random Assignment Problems," 2006. Unpublished mimeo, Harvard University.
Konishi, Hideo and Utku Ünver, "Games of Capacity Manipulation in the Hospital-Intern Market," *Social Choice and Welfare*, 2006, *27*, 3–24.
Ma, Jinpeng, "Stable Matchings and the Small Core in Nash Equilibrium in the College Admissions Problem," *Review of Economic Design*, 2002, *7*, 117–134.
McVitie, D. G. and L. B. Wilson, "Stable Marriage Assignments for Unequal Sets," *BIT*, 1970, *10*, 295–309.
Milgrom, Paul R., *Putting Auction Theory to Work*, Cambridge: Cambridge University Press, 2004.

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

Pathak, Parag A. and Tayfun Sönmez, "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 2008, *98*, 1636-1652.

Pesendorfer, Wolfgang and Jeroen Swinkels, "Efficiency and Information Aggregation in Auctions," *American Economic Review*, 2000, *90*, 499–525.

Romero-Medina, Antonio, "Implementation of Stable Solutions in a Restricted Matching Market," *Review of Economic Design*, 1998, *3*, 137–147.

Roth, Alvin E., "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 1982, *7*, 617–628.

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

– , "Misrepresentation and Stability in the Marriage Problem," *Journal of Economic Theory*, 1984b, *34*, 383–387.

– , "The College Admission Problem is not Equivalent to the Marriage Problem," *Journal of Economic Theory*, 1985, *36*, 277–288.

– , "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, Fisher-Schultz Lecture," *Econometrica*, 2002, *70*, 1341–1378.

– , "Deferred Acceptance Algorithms: History, Theory, Practice and Open Questions," *International Journal of Game Theory*, 36: 537-569.

– **and Elliot Peranson**, "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 1999, *89*, 748–780.

– **and John Vande Vate**, "Incentives in Two-sided Matching with Random Stable Mechanisms," *Economic Theory*, 1991, *1*, 31–44.

– **and Marilda A. O. Sotomayor**, *Two-sided Matching: a Study in Game-theoretic Modeling and Analysis*, Cambridge: Econometric Society monographs, 1990.

– **and Uriel Rothblum**, "Truncation Strategies in Matching Markets: In Search of Advice for Participants," *Econometrica*, 1999, *67*, 21–43.

– **, Tayfun Sönmez, and Utku Ünver**, "Kidney Exchange," *Quarterly Journal of Economics*, 2004, *119*, 457–488.

– **, – , and** – , "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review*, 2007, *97*, 828–851.

Rustichini, Aldo, Mark Satterthwaite, and Steven Williams, "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, 1994, *62*, 1041–1064.

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

– , "Games of Manipulation in Marriage Problems," *Games and Economic Behavior*, 1997b, *20*, 169–176.

– , "Can Pre-arranged Matches be Avoided in Two-Sided Matching Markets?," *Journal of Economic Theory*, 1999, *86*, 148–156.

Swinkels, Jeroen, "Efficiency of Large Private Value Auctions," *Econometrica*, 2001, *69*, 37–68.