Incentive Compatibility of Large Centralized Matching Markets

SangMok Lee *

(Job Market Paper)

First Version: July 3, 2011
Current Version: October 27, 2011

Abstract

This paper discusses the strategic manipulation of stable matching mechanisms. Stable matching mechanisms are very successful in practice, despite theoretical concerns that they are manipulable by participants. Our key finding is that most agents in large markets are close to being indifferent among partners in all stable matchings. It is known that the utility gain by manipulating a stable matching mechanism is bounded by the difference between utilities from the best and the worst stable matching partners. Thus, the main finding implies that the proportion of agents who may obtain a significant utility gain from manipulation vanishes in large markets. This result reconciles the success of stable mechanisms in practice with the theoretical concerns about strategic manipulation. We also introduce new techniques from the theory of random bipartite graphs for the analysis of large matching markets.

Keywords: Two-sided matching, Stable matching mechanism, Large market, Random bipartite graph

JEL Class: C78, D61, D78

*Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. E-mail: `sangmok-at-hss.caltech.edu` I am especially grateful to Federico Echenique, Matt Shum, and Leeat Yariv for encouragement and guidance. I also thank Hyoung-Jun Ahn, Itai Ashilagi, Luke Boosey, Kim Border, Chris Chambers, Guilherme Freitas, Cary Frydman, John Hatfield, Fuhito Kojima, Sera Linardi, Ruth Mendel, Juan Pereyra, Jean-Laurent Rosenthal, Tom Ruchti, Bob Sherman, Tayfun Sönmez, Bumin Yenmez, as well as seminar participants at SISL, the Midwest Economic Theory Meeting in Notre Dame, and the 2011 North American Summer Meeting and the Asian Meeting of the Econometric Society for very helpful discussions.

1 Introduction

1.1 Overview

In this paper, we study the most popular class of algorithms, called *stable matching mechanisms*, used in centralized matching markets, such as the National Resident Matching Program (NRMP) and School Choice Programs in NYC and Boston. A matching is regarded as stable if no agent is matched with an unacceptable partner, and there is no pair of agents on opposite sides of the market who prefer each other to their current partners. A stable matching mechanism takes preference reports by participants and produces a stable matching with respect to the submitted preferences. We ask how stable matching mechanisms remain so successful, despite the fact that the mechanisms are easily manipulable by the participants through misrepresenting their preferences. In particular, we analyze whether large markets, i.e. ones consisting of a large number of participants, would mitigate incentives to manipulate a stable matching mechanism.

Two-sided matching markets are markets with two kinds of agents, in which agents of one kind match with agents of the other kind. Examples of such markets include firms and workers in professional labor markets (Roth and Peranson, 1999), schools and students in school choice programs (Abdulkadiroglu and Sönmez, 2003), men and women in the marriage market or dating sites (Choo and Siow, 2006; Hitsch, Hortaçsu, and Ariely, 2010), birth mothers and potential adoptive parents in the market for child adoption (Bernal, Hu, Moriguchi, and Nagypal, 2007; Baccara, Collard-Wexler, Felli, and Yariv, 2010), and cadets and branches in the military (Sönmez and Switzer, 2011). Market designers seeking to achieve desirable outcomes to these matching markets have introduced centralized clearinghouses.

In market design, the concept of "stability" has been considered of central importance. In practice, successful mechanisms often implement a stable matching with respect to submitted preferences (Roth and Xing, 1994; Roth, 2002). The best-known market design examples, such as the NRMP and School Choice Programs in NYC and Boston, also use a particular stable matching mechanism, called the doctor-proposing or student-proposing Gale-Shapley algorithm.[1] Table 1 below lists whether each clearinghouse produces a stable matching with respect to submitted preferences, and whether these clearinghouses are still in use or no longer operating. With few exceptions, stable matching mechanisms have been successful for the most part whereas unstable mechanisms have mostly failed.[2]

From a theoretical perspective, however, stable matching mechanisms have a significant shortcoming. While the mechanisms produce stable matchings by assuming that all participants reveal their true preferences, in fact *no stable matching mechanism is strategy-proof* (Roth, 1982).[3] Par-

[1] The algorithm is customized for each application. For details of the actual algorithms applied, see Roth and Peranson (1999); Abdulkadiroglu, Pathak, and Roth (2009); Abdulkadiroglu, Pathak, Roth, and Sönmez (2006).

[2] Table 1 is reorganized from tables in Roth (2002) and McKinney, Niederle, and Roth (2003). The clearinghouse for the gastroenterology fellowship market is a rare case in which a stable matching mechanism started to fail in 1996, was abandoned in 2000, and then was reinstated in 2006 (Niederle and Roth, 2005; Roth, 2008).

[3] In fact, strategy-proofness is incompatible not only with stability but even with weaker conditions of Pareto efficiency and individual rationality (Alcalde and Barberà, 1994; Sönmez, 1999).

	Still in use	No longer in use
Stable	**The NRMP**: over 40 specialty markets and submarkets for first year postgraduate positions, and 15 for second year positions **Specialty matching services**: over 30 subspecialty markets for advanced medical residencies and fellowships **School choice programs**: NYC, Boston **Canadian lawyers:** multiple regions **British regional medical markets**: Edinburgh (≥‘69), Cardiff **Dental residencies**: 3 specialties **Other healthcare markets**: Osteopaths (≥‘94), Pharmacists, Clinical psychologists (≥‘99)	**Dental residencies**: Periodontists(<‘97), Prosthodontists (<‘00) **Canadian lawyers:** British Columbia(<‘96)
Unstable	**British regional medical markets**: Cambridge, London Hospital	**British regional medical markets**: Birmingham, Edinburgh (<‘67), Newcastle, Sheffield **Other healthcare markets**: Osteopaths (<‘94)

Table 1: Stable and unstable (centralized) mechanisms.

ticipants may achieve a better matching by misrepresenting their preferences, either by changing the order of the preference lists or by announcing that some acceptable agents are unacceptable. Even the NRMP and School Choice Programs in NYC and Boston, while widely acknowledged as a model of successful matching programs, cannot rule out such incentives for strategic misrepresentation. Indeed, the possibility of such manipulation is mostly unavoidable. Whenever there is more than one stable matching, at least one agent can profitably misrepresent her preferences (Roth and Sotomayor, 1990), and the conditions under which a preference profile contains a unique stable matching seem to be quite restrictive (Eeckhout, 2000; Clark, 2006).[4] Thus, markets are likely to have agents with an incentive to manipulate a stable matching mechanism. In addition, Pittel (1989) shows that the number of stable matchings tends to increase as the number of participants becomes large. Accordingly, when market designers deal with large markets, concerns regarding strategic manipulation are heightened. As stable matching mechanisms are not incentive compatible, the mechanisms may be manipulated by participants, thereby not implementing the intended matchings. Moreover, each participant's decision may become hard to make since she needs to best respond to other agents' strategic manipulations.

We consider matching markets that each firm hires one worker, a model which is known as a one-to-one matching. We measure incentives to manipulate a stable matching mechanism by assuming that each firm-worker pair receives utilities, one for the firm and the other for the worker, which in turn determine ordinal preferences. In order to study the likelihood of an agent having a significant incentive to manipulate, we assume that utilities are randomly drawn from some underlying distributions. The key finding of this paper is that the proportion of participants who can potentially achieve a significant utility gain from manipulation vanishes as the market becomes

[4] It is an open question to characterize the complete set of preference profiles containing a unique stable matching.

large. This result holds both when each agent knows the preferences of all other agents (complete information), and when an agent may not know the preferences of other agents (incomplete information). Given the tangible and intangible costs of strategic behavior in real life, we believe that this result may reconcile the success of stable matching mechanisms with the theoretical concerns about manipulability. In addition, based on this paper's finding, market designers may more confidently advise participants to submit their true preferences.

1.2 A Motivating Example

To understand the logic behind strategic manipulation, consider a simple labor market with three firms and three workers. We illustrate how, in such a situation, an agent can achieve a better partner by misrepresenting her preferences. In addition, we show that the best achievable partner from manipulation must be a partner in a stable matching under her true preferences.

Table 2 lists preferences of firms over workers, and of workers over firms which are known to all participants. For instance, firm 1 most prefers worker 3, followed by worker 1 and worker 2. Similarly, worker 1 most prefers firm 2, followed by firm 3 and firm 1. Under these preferences, there are two stable matchings: in one stable matching (marked by $\langle\cdot\rangle$), f_1, f_2, and f_3 are matched with w_1, w_2, and w_3, respectively; in the second stable matching (marked by $[\cdot]$), f_1, f_2, and f_3 are matched with w_2, w_1, and w_3, respectively.

$$\begin{array}{lcccccc} \mathbf{f_1} & : & w_3 & \succ & \langle w_1\rangle & \succ & [w_2] \\ \mathbf{f_2} & : & \langle w_2\rangle & \succ & [w_1] & \succ & w_3 \\ \mathbf{f_3} & : & \langle [w_3]\rangle & \succ & w_1 & \succ & w_2 \end{array} \quad , \quad \begin{array}{lcccccc} \mathbf{w_1} & : & [f_2] & \succ & f_3 & \succ & \langle f_1\rangle \\ \mathbf{w_2} & : & [f_1] & \succ & \langle f_2\rangle & \succ & f_3 \\ \mathbf{w_3} & : & f_2 & \succ & \langle [f_3]\rangle & \succ & f_1 \end{array}$$

Table 2: An example of a two-sided matching market with 3 firms and 3 workers.

Suppose that all agents submit their true preferences, and a stable matching mechanism produces the second stable matching marked by $[\cdot]$. In that case, suppose firm 1 misrepresent her preferences and announces that workers 3 and 1 are acceptable, but not worker 2. For the submitted preferences, there is a unique stable matching marked by $\langle\cdot\rangle$. The stable matching mechanism, which produces a stable matching for submitted preferences, will produce the matching marked by $\langle\cdot\rangle$. Ultimately, firm 1 is better off because firm 1 is matched with worker 1 rather than worker 2.

However, whichever preference list firm 1 submits, the firm will not be matched with worker 3. The pair (f_3, w_3) would otherwise block the matching. For instance, if f_1 declares that only w_3 is acceptable, then the only stable matching matches f_2 with w_2, and f_3 with w_3, and firm 1 will remain unmatched. More broadly, whenever a stable matching mechanism is applied, participants cannot be matched with a partner who is strictly preferred to all stable matching partners with respect to the initial preferences (Demange, Gale, and Sotomayor, 1987). Since participants are guaranteed to be matched with one of their stable matching partners, the gain from manipulation is bounded by the difference between the most and the least preferred stable matching partners. Based on the above observation, we mainly focus on the difference between the most and the least preferred stable matching partners.

1.3 Outline of the Paper

Prior to describing the model in detail, we briefly discuss the outline of the model, our main results, and the key idea behind the proof.

We consider a sequence of one-to-one matching markets, each of which has n firms and n workers. Preferences of firms over workers, or of workers over firms are generated by utilities, which are randomly drawn from some underlying distributions on $\mathbb{R}_+$.[5] We formulate utilities as the weighted sum of a *common-value* component and an *independent private-value* component. That is, when a firm f is matched with a worker w, the firm receives

$$U_{f,w} = \lambda\, U^o_w + (1-\lambda)\, \zeta_{f,w} \qquad (0 \leq \lambda \leq 1),$$

where U^o_w is the intrinsic value of w, which is common to all firms, and $\zeta_{f,w}$ is w's value as independently evaluated by firm f. In other words, any firm that is matched with a worker w receives the same common-value of the worker w, but receives distinct private-value of the worker w. We similarly define the utilities of workers.

The common-value component introduces a commonality of preferences, which is prevalent in real matching markets. In the entry-level labor market for doctors, for instance, the US News and World Report's annual rankings are often referred to as a guideline to the best hospitals. We also consider the pure private-value model ($\lambda = 0$) for theoretical reasons. In matching theory, commonality drives the uniqueness of stable matchings (Eeckhout, 2000; Clark, 2006), a situation in which no agent has an incentive to manipulate a stable matching mechanism (Roth and Sotomayor, 1990). If a preference profile has several stable matchings, commonality of preferences leads to smaller differences in utilities from stable matchings (Samet, 2011), so agents have less of an incentive to manipulate a stable matching mechanism. By including the pure private-value case in our model, we show that commonality may be beneficial, but is not necessary for incentive compatibility of stable matching mechanisms.

The main finding of the paper is that while agents in a large market typically have multiple stable partners, most agents are close to being indifferent among the all stable matching partners (Theorem 1).[6] We observed in the motivating example that when a stable matching mechanism is applied, the best an agent can achieve (by misrepresenting her preferences) is matching with her

[5] The only restrictions on distributions are bounded supports and some continuity conditions.

[6] The main theorem seems quite consistent with observations from real market applications. Pathak and Sonmez (2008) collect the data of students' preferences over schools in the new Boston school choice program, and show that the real market tends to have a very small number of stable matchings. (The preference data is reliable as truthfully revealing their preferences is a dominant strategy for students.) Both suggest that large matching markets tend to have small cores. In our theory of one-to-one matchings, we find small differences in utilities from stable matchings, whereas in the data from a many-to-one matching market, there is a small number of stable matchings.

best stable matching partner with regard to the true preferences (Demange, Gale, and Sotomayor, 1987). As such, our main finding implies that when a stable matching mechanism is applied and all other agents reveal their true preferences, the expected proportion of agents who have an incentive to manipulate the mechanism vanishes as the market becomes large.

Furthermore, we identify an ϵ-Nash equilibrium in which most participants report their true preferences.[7] In a large market, a small proportion of agents may still have large incentives to manipulate a stable matching mechanism. Under the identified equilibrium, we let those agents with significant incentives to manipulate do misrepresent their preferences. Nevertheless, the rest of participants still have no incentive to respond to such manipulations. More precisely, we show that for any $\epsilon > 0$ with high probability a large market has an ϵ-Nash equilibrium in which most participants reveal their true preferences (Corollary 2).

From a methodological standpoint, our paper is the first to introduce techniques from *random bipartite graph theory* to matching models. To prove the main theorem, we basically need to *count* the number of firms and workers satisfying certain conditions. The theory of random bipartite graphs provides techniques to count the likely numbers of firms and workers satisfying the specified conditions. More precisely, we draw a graph with a set of firms and workers whose common-values are above certain levels. We join each firm-worker pair by an edge if one of their independent private-values is significantly lower than the upper bound of the support. It turns out that every firm-worker pair where both the firm and the worker fail to achieve certain threshold levels of utility in a stable matching must be joined by an edge. Their private-values would otherwise both be so high that they would prefer each other to their current partners, and thus block the stable matching. For each realized graph, we consider the bi-partitioned subset of nodes, i.e. firms and workers, such that every pair of nodes, one from each partition, is joined by an edge. It is known that the possibility of having a relatively large such subset of nodes ultimately becomes infinitesimal as the initial set of nodes becomes large (Dawande, Keskinocak, Swaminathan, and Tayur, 2001). That is, in terms of the matching model, the set of firms and workers, whose common-values are high but who fail to achieve high levels of utility, will remain relatively small as the market becomes large.

This paper mainly focuses on the case of complete information, in which all participants are aware of all other agents' preferences. Nevertheless, we can extrapolate its findings to a market with incomplete information, in which each agent is partially informed about other agents' preferences. Various setups are conceivable: an agent may know (i) only her own utilities from matching with agents on the other side; (ii) her own utilities and common-values from matching with agents on the other side; (iii) her own utilities, common-values from matching with agents on the other side, and her own common-value to agents matching with her; or (iv) her own utilities and all agents' common-values. Regardless of the information structure, the key finding from the complete information case still holds with incomplete information. That is, most agents are ex-ante close

[7]Under an ϵ-Nash equilibrium, agents are approximately best responding to other agents' strategies such that no one can gain more than ϵ by switching to an alternative strategy.

to being indifferent among all stable matchings in a large market (Theorem 4). This is because with high probability agents are close to being indifferent among realized stable matching partners, which is this study's key finding in the context of complete information.

However, we do not find an equilibrium corresponding to the ϵ-Nash equilibrium under complete information. If some agents manipulate a stable matching mechanism based on the expected utility gain from manipulation, they may become worse off afterwards. This may, in turn, expand the differences between utilities generated by stable matchings to other agents. Consequently, other agents may misrepresent their preferences as a best response to the manipulation.

1.4 Related Literature

Strategic manipulability has been a major concern in market design. Hence, a number of studies have addressed the incentives to manipulate a stable matching mechanism (Roth and Peranson, 1999; Immorlica and Mahdian, 2005; Kojima and Pathak, 2009). These studies consider a particular stable matching mechanism, the worker-proposing Gale-Shapley algorithm, which implements a stable matching favorable to workers. As truthfully revealing their preferences is a dominant strategy for workers in this mechanism (Roth, 1982; Dubins and Freedman, 1981), the papers focus on firms' incentives to misrepresent their preferences.

Unlike the current paper, these studies assume that firms will manipulate a mechanism regardless of how much benefit the firms can obtain by so doing. In particular, a firm has no incentive to misrepresent its preferences if and only if it has a unique stable matching partner (Roth and Sotomayor, 1990). Thus, the primary goal is to find conditions on a preference profile in which most firms have a unique stable matching partner. As Roth and Peranson (1999) also point out, a crucial assumption is that agents on one side (say workers) consider only up to a fixed number of agents on the other side acceptable, even when the market size has become large. Under this assumption, Roth and Peranson, based on a computational analysis, show that the proportion of firms who have more than one stable matching partner converges to zero as the market becomes large. This convergence is theoretically proven by Immorlica and Mahdian and extended to the case of many-to-one matchings by Kojima and Pathak.

The main advantage of our approach is that we obtain non-manipulability of stable matching mechanisms as a pure property of market size, without resorting to the assumption of limited acceptability. In fact, the assumption of limited acceptability may leads to large market models that do not match basic features of real applications. Even with a weak commonality of preferences, the proportion of firms who are accepted by at least some workers may become small as the market becomes large. In this case, most firms do have a unique stable matching partner, but quite often the unique stable matching partner is only the firm itself: i.e. *a large proportion of agents remain unmatched.*

Figure 1 presents this phenomenon with simulations in which each worker considers only up to 30 most preferred firms acceptable. The utility of a firm is defined as $U_{f,w} = \lambda\, U^o_w + (1-\lambda)\, \zeta_{f,w}$, and the utility of a worker is similarly defined. The value of each component is drawn from the uniform distribution over [0, 1]. Each graph depicts the proportion of firms (or workers) unmatched in stable matchings averaged over 10 repetitions.[8] Even with modest levels of commonality of preferences, the proportion of unmatched agents in stable matchings increases as the market becomes large. It is worth noting that these simulations are based on preferences generated, not by the previous studies' model, but by our own. Thus, the simulations do not directly represent features of the previous studies. However, we observe the similar effects of the limited acceptability assumption in simulations based on the previous studies' model. We provide additional simulation results in Appendix E.

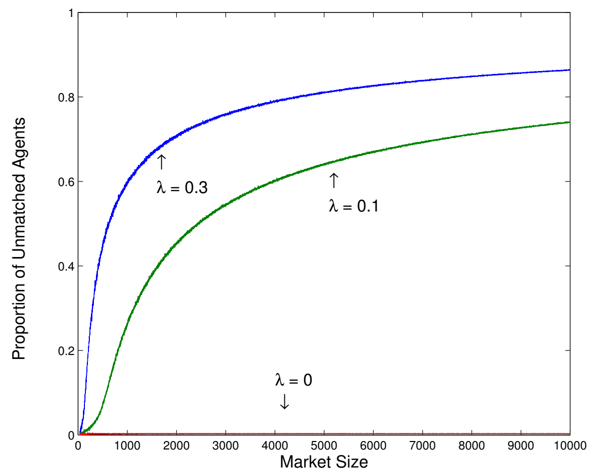


Figure 1: Proportion of agents unmatched in stable matchings.

Another strand of literature on large matching markets considers a market where a finite number of firms are matched with a continuum of workers (Azevedo and Leshno, 2011). It is shown that generically each market has a unique stable matching, to which the set of stable matchings in markets with large discrete workers converges. Based on this model, Azevedo (2010) studies firms' incentives to manipulate capacities to hire workers. The paper also compares welfare effects between situations where each firm pays its employees equally (uniform wages) and those where each firm may pay different wages to different workers (personalized wages). While previous studies with fixed capacities suggest that a uniform wage may induce inefficient matching and compress workers' wages (Bulow and Levin, 2006; Crawford, 2008), if firms can manipulate their capacities, the uniform wage may produce higher welfare as they cause less capacity reduction.

The large market approach is not limited to the standard matching model. Ashlagi, Braverman, and Hassidim (2011) and Kojima, Pathak, and Roth (2010), for instance, develop models of large matching markets with couples. When couples are present, notwithstanding the concerns about strategic manipulation, a market does not necessarily have a stable matching (Roth, 1984). These

[8] Given a preference profile, the set of unmatched agents is the same for all stable matchings (McVitie and Wilson, 1970).

studies show that the probability that a market with couples contains a stable matching converges to one as the market becomes large. Moreover, when a mechanism produces a stable matching with high probability, it is an approximate equilibrium for all participants to submit their true preferences. The results are based on the condition that the number of couples grows slower than the market size, with some additional regularity conditions.[9]

In the assignment problem of allocating a set of indivisible objects to agents, Kojima and Manea (2010) study incentives in the probabilistic serial mechanism (Bogomolnaia and Moulin, 2001). The probabilistic serial mechanism is proposed as a mechanism that improves the ex-ante efficiency of the random priority mechanism: All agents have higher chances of obtaining more preferred objects by using the probabilistic serial mechanism. However, while the random priority mechanism is strategy-proof, the probabilistic serial mechanism is not. Kojima and Manea show that for a fixed set of object types and an agent with a given utility function, if there is a sufficiently large number of copies of each object type, then reporting true preferences is a weakly dominant strategy for the agent.[10]

The rest of this paper is organized as follows. In Section 2, we introduce our model – a sequence of matching markets with random utilities. In Section 3, we state the main theorem informally and then formally, and find an equilibrium behavior which may reconcile the conflicting features of stable matching mechanisms. In Section 4, we illustrate the intuition of the proof using a random bipartite graph model. In Section 5, we study a market with incomplete information. The conclusion of the paper is provided in Section 6. All detailed proofs and simulation results are relegated to the Appendix, which also includes definitions and related theorems of asymptotic statistics.

2 Model

The model is based on the standard one-to-one matching model. We introduce latent utilities, which in turn generate ordinal preferences.

2.1 Standard Two-sided Matching Model (Roth and Sotomayor (1990))

There are n firms and an equal number of workers. We denote the set of firms by F and the set of workers by W. Each firm has a strict preference list $\succ_f$ such as

$$\succ_f = w_1, w_2, w_3, f, \ldots, w_4.$$

[9] Ashlagi, Braverman, and Hassidim (2011) considers a market where the number of positions offered by firms exceeds the number of workers. Kojima, Pathak, and Roth (2010) inherits the assumption from Kojima and Pathak (2009) that agents on one side consider only up to a fixed number of agents on the other side acceptable.

[10] Che and Kojima (2010) show that the random assignments in the two mechanisms converge to each other as the number of copies of each object type goes to infinity. More generally, Liu and Pycia (2011) show that, including the two mechanisms, all sensible and asymptotically symmetric, strategy-proof, and ordinal efficient allocation mechanisms coincide asymptotically.

This preference list indicates that w_1 is firm f's first choice, w_2 is the second choice, and that w_3 is the least preferred worker that the firm still wants to hire. We also write $w \succ_f w'$ to mean that f prefers w to w'. We call a worker w **acceptable** to f if $w \succ_f f$, otherwise we call the worker **unacceptable**. We define $\succ_w$ similarly for each $w \in W$, and call $\succ := ((\succ_f)_{f\in F}, (\succ_w)_{w\in W})$ **a preference profile**.

A **matching** μ is a function from the set $F \cup W$ onto itself such that (i) $\mu^2(x) = x$, (ii) if $\mu(f) \neq f$ then $\mu(f) \in W$, and (iii) if $\mu(w) \neq w$ then $\mu(w) \in F$. We say a matching μ is **individually rational** if each firm or worker is matched to an acceptable partner, or otherwise remains unmatched. For a given matching μ, a pair (f, w) is called a **blocking pair** if $w \succ_f \mu(f)$ and $f \succ_w \mu(w)$. We say a matching is **stable** if it is individually rational and has no blocking pair.

For two stable matchings μ and μ', we write $\mu \succeq_i \mu'$ if an agent i weakly prefers μ to μ': i.e. $\mu(i) \succ_i \mu'(i)$ or $\mu(i) = \mu'(i)$. We also write $\mu \succeq_F \mu'$ if every firm weakly prefers μ to μ': i.e $\mu(f) \succeq_f \mu'(f)$ for every $f \in F$. Similarly, we write $\mu \succeq_W \mu'$ if every worker weakly prefers μ to μ': i.e. $\mu(w) \succeq_w \mu'(w)$ for every $w \in W$. A stable matching μ_F is **firm-optimal** if every firm weakly prefers it to any other stable matching μ: i.e. $\mu_F \succeq_F \mu$. Similarly, a stable matching μ_W is **worker-optimal** if every worker weakly prefers it to any other stable matching μ: i.e. $\mu_W \succeq_W \mu$. It is known that every market instance has a firm-optimal stable matching μ_F and a worker-optimal stable matching μ_W (Gale and Shapley, 1962): i.e. for any stable matching μ, we have $\mu_F \succeq_F \mu$ and $\mu_W \succeq_W \mu$. Moreover if μ and μ' are both stable matchings, then $\mu \succeq_F \mu'$ if and only if $\mu' \succeq_W \mu$ (Knuth, 1976). Thus for any stable matching μ, it must be the case that $\mu \succeq_F \mu_W$ and $\mu \succeq_W \mu_F$.

With some abuse of notation, we let μ denote a function $\succ \longmapsto \mu(\succ)$ from the set of all preference profiles to the set of all matchings. We call the function μ a **matching mechanism**, and say that a mechanism μ is **stable** if $\mu(\succ)$ is a stable matching with respect to preference profile $\succ$. We also let μ_F and μ_W denote firm-optimal and worker-optimal stable matching mechanisms. A matching mechanism induces a game in which each agent $i \in F \cup W$ states her preference list $\succ_i$. If for all $\succ_i$ and $\succ_{-i}$,

$$\mu(\succ_i^*, \succ_{-i}) \succeq_i \mu(\succ_i, \succ_{-i}),$$

then we call $\succ_i^*$ a **dominant strategy** for the agent i. A mechanism μ is called **strategy-proof** if it is a dominant strategy for every agent to state her true preference list.

2.2 Random Utilities

In order to measure incentives to manipulate a stable matching mechanism, we assume that preferences are induced by underlying utilities. Moreover, in order to measure likely incentives, we assume that the utilities are drawn from some underlying probability distributions.

We represent utilities by $n \times n$ random matrices $U = [U_{f,w}]$ and $V = [V_{f,w}]$. When a firm f and a worker w match with one another, the firm f receives utility $U_{f,w}$ and the worker w receives utility $V_{f,w}$. We let u and v denote realized matrices of U and V. For each pair (f, w), utilities are defined as

$$U_{f,w} = \lambda\, U^o_w + (1-\lambda)\, \zeta_{f,w} \qquad \text{and}$$
$$V_{f,w} = \lambda\, V^o_f + (1-\lambda)\, \eta_{f,w} \qquad (0 \le \lambda \le 1).$$

We call U^o_w and V^o_f *common-values*, and $\zeta_{f,w}$ and $\eta_{f,w}$ *independent private-values.*

Common-values are defined as random vectors

$$U^o := \langle U^o_w \rangle_{w\in W} \quad \text{and} \quad V^o := \langle V^o_f \rangle_{f\in F}.$$

Each U^o_w and V^o_f are drawn from distributions with positive density functions and with bounded supports in $\mathbb{R}_+$.[11] Independent private-values are defined as $n \times n$ random matrices

$$\zeta := [\zeta_{f,w}] \quad \text{and} \quad \eta := [\eta_{f,w}].$$

Each $\zeta_{f,w}$ and $\eta_{f,w}$ are randomly drawn from continuous distributions with bounded supports in $\mathbb{R}_+$. We assume that the utility of remaining unmatched is equal to 0.[12]

A random market is defined as a tuple $\langle F, W, U, V\rangle$, and a market instance is denoted by $\langle F, W, u, v\rangle$. Each firm f receives distinct utilities from different workers with probability 1. Thus for each $\langle F, W, u, v\rangle$, we can derive a strict preference list $\succ_f$ as

$$\succ_f = w, w', \ldots, w''$$

if and only if

$$u_{f,w} > u_{f,w'} > \cdots > u_{f,w''}.$$

We study properties of stable matchings in a sequence of random markets $\langle F_n, W_n, U_n, V_n\rangle_{n=1}^{\infty}$. The index n will be omitted whenever to do so does not lead to confusion.

The model includes both cases of a commonality of preferences ($\lambda > 0$) and pure private-values ($\lambda = 0$). The common-values introduce a commonality of preferences among firms over workers, and among workers over firms. When $\lambda > 0$, firms with high level of common-values tend to be ranked higher by workers, and vice versa. If $\lambda = 0$, all utilities are i.i.d, so a firm's ordering of workers are equally likely to be any permutation from the set of all permutations of n workers. Similarly, a worker's ordering of firms are equally likely to be any permutation from the set of all permutations of n firms.

In practice, commonality of preferences is prevalent. In the NRMP, some hospitals are consid-

[11] The bounded support condition of common-values is not necessary for the results with complete information (Theorem 1 and Corollary 2), and thus is not used in the proofs in Section C. We use the condition later for results with incomplete information (Theorem 4). When the supports are not bounded, the difference in the utilities of a firm from its most preferred workers tends to increase as the market becomes large.

[12] In terms of preferences induced by utilities, this assumption implies that all workers are acceptable to firms, and all firms are acceptable to workers. We impose this condition to keep the model tractable. If we relax the assumption, however, we obtain even stronger incentive compatibility, as far as all firms and workers are matched in stable matchings under true preferences. The intuition is similar to that behind the proof of Corollary 2.

ered prestigious and some doctors are considered very well-qualified. The common-value component provides a way of taking into account such commonality of preferences, while retaining the tractability of the model.

Although the pure private-value case ($\lambda = 0$) hardly represents any real application, it is theoretically valuable to include it in our model. Commonality drives the uniqueness of stable matchings (Eeckhout, 2000; Clark, 2006), a condition in which no agent has an incentive to misrepresent her preferences in a stable matching mechanism (Roth and Sotomayor, 1990). Samet (2011) also proposes commonality as a source establishing a small core: the small difference between utilities from the stable matchings favorable to firms, and to workers. By including the pure private-value case in our model, we can highlight that non-manipulability of stable matching mechanisms is a property solely derived from market size. Commonality may contribute to, but is not necessary for, incentive compatibility of stable matching mechanisms.[13]

3 Main Results

We informally state the main theorem, and then restate it with formal expressions. Later, we find an equilibrium behavior of a game induced by a stable matching mechanism in which most agents reveal their true preferences.

3.1 Stable Matchings in Large Markets

We first show that, while agents in a large market typically have multiple stable matching partners, most agents are close to being indifferent among the stable matching partners.

Theorem *For every $\epsilon > 0$, the expected proportion of firms (and workers) whose utility from one stable matching is within ϵ of the utility from any other stable matching, converges to one as the market becomes large.*

Corollary *For any positive cost of misrepresenting preferences, if other agents truthfully reveal their preferences, the expected proportion of agents who have no incentive to manipulate a stable matching mechanism converges to one as the market becomes large.*

It has been known that no stable matching mechanism is strategy-proof (Roth, 1982). For instance, when the worker-optimal matching mechanism (e.g. worker-proposing Gale-Shapley algorithm) is applied, although it is a dominant strategy for every worker to state her true preference list (Roth, 1982; Dubins and Freedman, 1981), there might be a firm which can become better off by misrepresenting its preference list. Noting that a matching mechanism is defined over all possible preference profiles, we may expect that a stable matching mechanism is not manipulable in

[13] When preferences have a strong commonality, a stable matching mechanism may have a higher chance to fail by unraveling instead of strategic preference misrepresentation (Halaburda, 2010). In any case, our model includes all degrees of commonality of preferences.

most cases of preference profiles. Unfortunately, though, it turns out that whenever there is more than one stable matching, at least one agent can profitably misrepresent her preferences (Roth and Sotomayor, 1990), and the condition of a preference profile containing a unique stable matching seems to be quite restrictive (Eeckhout, 2000; Clark, 2006).

However, the gain by manipulation is bounded even when agents form a coalition and coordinate the members' strategic behavior. Not all firms in the coalition will prefer the new matching outcome to the firm-optimal stable matching with respect to the true preferences, and not all workers in the coalition will prefer the new matching outcome to the worker-optimal stable matching with respect to the true preferences (Demange, Gale, and Sotomayor, 1987). Formally, let $\succ$ be a true preference profile, and let $\succ'$ differ from $\succ$ in that some coalition S of firms and workers misstate their preferences. Then, there is no matching, stable under $\succ'$, which is strictly preferred to every stable matching under $\succ$ by all members of S. If a coalition consists of a single firm, then the best the firm can achieve is matching with the firm-optimal stable matching partner with respect to the true preferences. Likewise, the best a worker can achieve is matching with the worker-optimal stable matching partner. Since every firm and worker is guaranteed to be matched with a stable matching partner without any strategic manipulation, the gain by manipulation is bounded by the difference between utilities from the firm-optimal and the worker-optimal stable matching partners.

As such, the main theorem implies that agents in a large market are most likely to have only a slight utility gain by misrepresenting their preferences, given that all other agents reveal their true preferences. For any given cost of misrepresenting preferences, if a market is large, participants are most likely to find no incentive to manipulate a stable matching mechanism.

In order to see whether a real market is large enough to mitigate incentives to manipulate stable matching mechanisms, we simulate our model with a market size of 26,000, roughly the same size of the NRMP in 2011.[14] We generate firms' and workers' utilities from common-values and independent private-values, each of which is randomly drawn from the uniform distribution over $[0, 1]$. Table 3 presents the proportion of firms whose differences in utilities generated by stable matchings are less than 0.05 (upper table) and 0.01 (lower table). The results show that for reasonable degrees of commonality of preferences, the size of the NRMP is large enough such that most agents would not have a significant incentive to manipulate a stable matching mechanism.

Formal Statement Given a market instance $\langle F, W, u, v\rangle$ and a matching μ, we let $u_\mu(\cdot)$ and $v_\mu(\cdot)$ denote utilities from the matching outcome: i.e. $u_\mu(f) := u_{f,\mu(f)}$ and $v_\mu(w) := v_{\mu(w),w}$. For each $f \in F$, we define $\Delta(f; u, v)$ as the difference between utilities from firm-optimal and worker-optimal stable matching outcomes: i.e.

$$\Delta(f; u, v) := u_{\mu_F}(f) - u_{\mu_W}(f).$$

[14] In 2011, there were 30,589 active applicants and 26,158 positions offered by 4,235 programs. See http://www.nrmp.org/data/resultsanddata2011.pdf and http://www.nrmp.org/res_match/about_res/impact.html.

λ	0.2	0.4	0.6	0.8
Result 1	97.41%	98.83%	99.39%	99.93%
Result 2	97.44%	98.79%	99.42%	99.92%
Result 3	97.43%	98.67%	99.47%	99.95%

(Differences in utilities < 0.05)

λ	0.2	0.4	0.6	0.8
Result 1	92.84%	96.64%	98.00%	99.44%
Result 2	93.04%	96.70%	98.10%	99.32%
Result 3	92.91%	96.52%	98.28%	99.48%

(Differences in utilities < 0.01)

Table 3: Proportions of firms with small differences in utilities (n=26,000)

Then, for every $\epsilon > 0$, we have the set of firms whose utilities are within ϵ of one another for all stable matchings, which is denoted by

$$A^F(\epsilon; u, v) := \{f \in F \mid \Delta(f; u, v) < \epsilon\}.$$

The previous theorem is an informal statement of the following theorem. We have similar notations and a theorem for workers, which are omitted here.

Theorem 1. *For every $\epsilon > 0$,*

$$E\left[\frac{|A^F(\epsilon; U, V)|}{n}\right] \to 1, \quad as \quad n \to \infty.$$

3.2 Equilibrium Analysis

Previously, we showed that most agents have no incentive to manipulate a stable matching mechanism as a market becomes large. However, the result requires the condition that all other participants reveal their true preferences. This condition is problematic since a small proportion of agents may still have large incentives to misrepresent their preferences. We may want to derive incentive compatibility as equilibrium behavior of a game induced by a stable matching mechanism.

In fact, the main theorem implies that with high probability a large market has a natural equilibrium in which most agents reveal their true preferences. We first state this finding as a corollary, and then describe appealing aspects of the equilibrium behavior and the intuition behind the proof.

Corollary 2. *For any $\epsilon, \delta, \theta > 0$, there exists N such that with probability at least $(1-\delta)$ a market of size $n > N$ has an ϵ-Nash equilibrium in which $(1-\theta)$ proportion of agents reveal their true preferences.*

This corollary is based on *simple* equilibrium behavior. Most agents simply reveal their true preferences. Agents misrepresenting their preferences use *truncation strategies*: an agent submits a preference list of the first k ($k < n$) in the same order as her true preference list. Truncations are natural strategies. Agents do not need to carefully devise the order of the preference list. In addition, truncation strategies are undominated, or, in other words, have "a best response property" (Roth and Vande Vate, 1991). If a stable matching mechanism is applied, for any given submitted preferences by other agents, an agent always has a best response that is a truncation of her true preference list.[15]

For each market instance $\langle F, W, u, v\rangle$, we consider an ϵ-Nash equilibrium in which some (not necessarily all) agents, who have potential gains from manipulations larger than ϵ, submit truncations of their true preferences. If there exists a stable matching under the true preferences remaining individually rational under the announced preferences, then for all participants the difference between utilities from firm-optimal and worker-optimal stable matchings decreases. Specifically, let $\succ$ be a true preference profile and $\succ'$ differ from $\succ$ in that some coalition of firms and workers misstate their preferences using truncations. If there exists at least one matching μ stable under $\succ$ remaining individually rational under $\succ'$, then all stable matchings for $\succ'$ are also stable under $\succ$. Thus, truncations by some agents result in smaller differences in utilities from stable matchings for all participants.

This property follows because truncations do not create additional blocking pairs. If a matching μ, which is stable under $\succ$, remains individually rational under $\succ'$, then μ is indeed stable under $\succ'$ since no blocking pair has been generated by truncations. Noting that the set of unmatched agents is the same for all stable matchings (McVitie and Wilson (1970)), all participants are matched in stable matchings under $\succ'$.[16] Then, any stable matching μ' with regard to $\succ'$ is also stable under $\succ$. If (f, w) is a blocking pair of μ' with respect to $\succ$, then it would have been a blocking pair of μ' with respect to $\succ'$, which contradicts that μ' is stable under $\succ'$.

For any preference profile and for any coalition of participants, there exist truncations by members of the coalition such that at least one stable matching under true preferences remains individually rational, and those who truncate their preferences have no incentive to truncate further. Then, participants who initially have smaller than ϵ differences in utilities from stable matchings will have even less differences in utilities from stable matchings under the announced preferences. Thus, these participants have no incentive to respond to others' truncations, thereby submitting their true preferences. Lastly, Theorem 1 guarantees that most participants are the ones revealing their true preferences.[17]

[15] Furthermore, when agents do not have complete information about the preference profile, truncation strategies require less information to manipulate a stable mechanism (Roth and Rothblum, 1999).

[16] Here, we use the condition that all participants are matched in stable matchings under $\succ$. If some agents are unmatched in stable matchings due to, for instance, unequal populations or unacceptable agents, we need an additional condition that agents would truncate their preferences only when truncations are strictly profitable. In particular, if an agent is unmatched in stable matchings under $\succ$, the agent will remain unmatched when she truncates her preference list. If these unmatched agents do not truncate their preference lists, then we obtain the same result: all stable matchings under $\succ'$ are stable under $\succ$, provided that there exists a stable matching under $\succ$ remaining individually rational under $\succ'$. The proof is easy to derive, and thus we omit it here.

[17] We use an equivalent statement of Theorem 1. Note that $|A^F(\epsilon; U, V)|/n$ is bounded above by 1 with probability 1. By using Theorem A.1 and Theorem A.2, we shall rewrite Theorem 1, written as convergence in mean, as the

4 Intuition Behind the Proof of Theorem 1

To prove Theorem 1, we take distinct approaches for the pure common-value case ($\lambda = 1$), the pure private-value case ($\lambda = 0$), and the general cases ($0 < \lambda < 1$).

For the pure common-value case ($\lambda = 1$), there exists a unique stable matching, so the theorem follows immediately. A stable matching sorts firms and workers such that a firm and a worker in the same rank will be matched with one another. Consider the firm-worker pair with the highest common-values. The pair must be matched in a stable matching. If it were otherwise, the firm would prefer the worker to its partner and the worker would prefer the firm to her partner, and thus they would form a blocking pair. By sequentially applying the same argument to pairs with the next highest common-values, we find that assortative matching is a unique stable matching.

For the pure private-value case ($\lambda = 0$), we still derive the theorem relatively easily from Pittel (1989). Pittel considers a model that is essentially the same as our pure private-value model ($\lambda = 0$), and analyzes the sum of each firm's partner's rank number in the worker-optimal stable matching.[18] When each firm ranks workers in order of preferences (i.e. the most preferred worker is ranked 1, the next worker is ranked 2, and so on), Pittel shows that the sum of the rank numbers of firms' partners in the worker-optimal stable matching is asymptotically equal to $n^2 \log^{-1} n$. Then, the rank number of each firm is roughly $n \log^{-1} n$ on average. In turn, as we normalize the rank number by the market size n, the normalized average rank number is roughly equal to $\log^{-1} n$, converging to 0. As the private-values are randomly drawn from distributions with bounded supports, even the worst stable matching gives utilities asymptotically close to the upper bound. Therefore, all stable matchings yield only slightly different utilities.

For the general cases ($0 < \lambda < 1$), however, the probability distribution over preference profiles becomes complicated and intractable. Accordingly, we directly analyze the asymptotic utilities rather than referring to the corresponding preference rank numbers. Basically, we want to count participants whose utilities from all stable matchings are slightly different from each other. We therefore need techniques of counting for which we use the bipartite graph theory. We interpret the set of firms and workers as a bi-partitioned set of nodes and draw a graph based on the realized utilities. Then, since the utilities are random, the theory of random bipartite graphs provides us with techniques to count the likely numbers of nodes, i.e. firms and workers, meeting specified conditions. Since the theory of random bipartite graphs has not been used before in the matching literature, we describe the techniques in greater depth in the following subsection.

We relegate detailed proofs for the cases of $\lambda = 0$ and $0 < \lambda < 1$ to Appendix B and Appendix C, respectively.

---

following convergence in probability: for any $\epsilon, \delta, \theta > 0$, there exists N such that

$$P\left(\frac{|A^F(\epsilon; U, V)|}{n} > 1 - \theta\right) > 1 - \delta, \quad \text{for every} \quad n > N.$$

[18] Pittel does not consider utilities, but a model with random preference profiles. As all preference profiles are equally likely to occur, though, the model is essentially the same as our pure private-value model ($\lambda = 0$).

4.1 A Random Bipartite Graph Model

A **graph** G is a pair (V, E), where V is a set called **nodes** and E is a set of unordered pairs (i, j) or (j, i) of $i, j \in V$ called **edges**. The nodes i and j are called the **endpoints** of (i, j). We say that a graph $G = (V, E)$ is **bipartite** if its node set V can be partitioned into two disjoint subsets V_1 and V_2 such that each of its edges has one endpoint in V_1 and the other in V_2. A **biclique** of a bipartite graph $G = (V_1 \cup V_2, E)$ is a set of nodes $U_1 \cup U_2$ such that $U_1 \subset V_1$, $U_2 \subset V_2$, and for all $u_1 \in U_1$ and $u_2 \in U_2$, $(u_1, u_2) \in E$. In other words, a biclique is a complete bipartite subgraph of G. We say that a biclique is **balanced** if $|U_1| = |U_2|$, and refer to a balanced biclique with the maximum number of nodes as **a maximum balanced biclique**.

Given a partitioned set $V_1 \cup V_2$, we consider a random bipartite graph $G(V_1 \cup V_2, p)$. A bipartite graph $G = (V_1 \cup V_2, E)$ is constructed so that each pair of nodes, one in V_1 and the other in V_2, is included in E independently with probability p. We use the following theorem in the proof.

Theorem 3 (Dawande, Keskinocak, Swaminathan, and Tayur (2001))**.** *Consider a random bipartite graph $G(V_1 \cup V_2, p)$, where $0 < p < 1$ is a constant, $|V_1| = |V_2| = n$, and $\beta(n) = \log n / \log \frac{1}{p}$. If the maximum balanced biclique of this graph has size $B \times B$, then*

$$P\left(\beta(n) \leq B \leq 2\beta(n)\right) \to 1, \quad \textit{as} \quad n \to \infty.$$

4.2 Intuition of the Proof ($0 < \lambda < 1$)

Roughly stated, we observe that stable matchings become assortative-like matchings as a market becomes large: firms with higher common-values become more likely to match with workers with higher common-values. We illustrate this assortative-like feature of stable matchings by introducing a 3-tier market. In a 3-tier market, firms and workers are partitioned into three tiers, and endowed with tier-specific common-values. Then, most firms and workers in the same tier are matched with each other in assortative-like stable matchings. In this situation, the expected proportion of firms in tier-1, which fail to achieve high levels of utility converges to 0 as the market becomes large. We demonstrate how to use techniques from the theory of random bipartite graphs as we prove this observation formally.

In a 3-tier market, F is partitioned into F_1, F_2, and F_3; and W is partitioned into W_1, W_2, and W_3. For simplicity, we assume that all tiers are of equal size:

$$|F_k| = |W_k| = n/3 \qquad (k = 1, 2, 3).$$

If $f \in F_k$ and $w \in W_l$ are matched with one another, then they receive utilities

$$U_{f,w} = u_l^o + \zeta_{f,w} \quad \text{and} \quad V_{f,w} = v_k^o + \eta_{f,w}.$$

Common-values are uniquely determined by tiers such that

$$u_1^o > u_2^o > u_3^o \quad \text{and} \quad v_1^o > v_2^o > v_3^o.$$

Private-values, $\zeta_{f,w}$ and $\eta_{f,w}$, are randomly drawn from uniform distributions over $[0, \bar{u}]$ and $[0, \bar{v}]$, respectively. In other words, the firm receives tier-specific common-value corresponding to the worker's tier added to independent private-value, and the worker receives tier-specific common-value corresponding to the firm's tier added to independent private-value. We, without loss of generality, ignore λ and $(1-\lambda)$ by incorporating the weights into the tier-specific common-values and the distributions of independent private-values.

We find an asymptotic lower bound on utilities that tier-1 firms receive in a stable matching mechanism. The lower bound is defined as the level arbitrarily close to the maximal utility that a firm can achieve by matching with tier-2 workers: i.e. $u_2^o + \bar{u} - \epsilon$. That is, firms in tier-1 achieve high levels of utility by levering on the existence of tier-2 workers. Although not necessarily being matched with tier-2 workers, firms in tier-1 would otherwise make blocking pairs with workers in tier-2. Formally, we define the set of tier-1 firms that fail to achieve the specified utility level in the worker-optimal stable matching as

$$\bar{F} := \{f \in F_1 \mid u_{\mu_W}(f) \leq u_2^o + \bar{u} - \epsilon\},$$

and show that

$$E\left[\frac{|\bar{F}|}{n/3}\right] \to 0, \quad \text{as} \quad n \to \infty.$$

Given realized private-values, we draw a bipartite graph with the set of firms in tier-1, and workers in tiers up to 2 (i.e. tier-1 and tier-2) as a bi-partitioned set of nodes (see the left figure in Figure 2). Each pair of $f \in F_1$ and $w \in W_1 \cup W_2$ is joined by an edge if and only if one of their private-values is low:

$$\zeta_{f,w} \leq \bar{u} - \epsilon \quad \text{or} \quad \eta_{f,w} \leq \bar{v} - (v_1^o - v_2^o).$$

We define the set of workers in tiers up to 2 matched with non tier-1 firms as

$$\bar{W} := \{w \in W_1 \cup W_2 \mid \mu_W(w) \notin F_1\}.$$

Then, $\bar{F} \cup \bar{W}$ is a biclique: i.e. every firm-worker pair from $\bar{F}$ and $\bar{W}$ is joined by an edge (as illustrated by the right figure in Figure 2).

To see why $\bar{F} \cup \bar{W}$ is a biclique, suppose that $f \in \bar{F}$ and $w \in \bar{W}$ are not joined. Since $f \in \bar{F}$,

$$u_{\mu_W}(f) \leq u_2^o + \bar{u} - \epsilon.$$

Since $w \in \bar{W}$, the worker is not matched with a tier-1 firm, and thus

$$v_{\mu_W}(w) \leq v_2^o + \bar{v}.$$

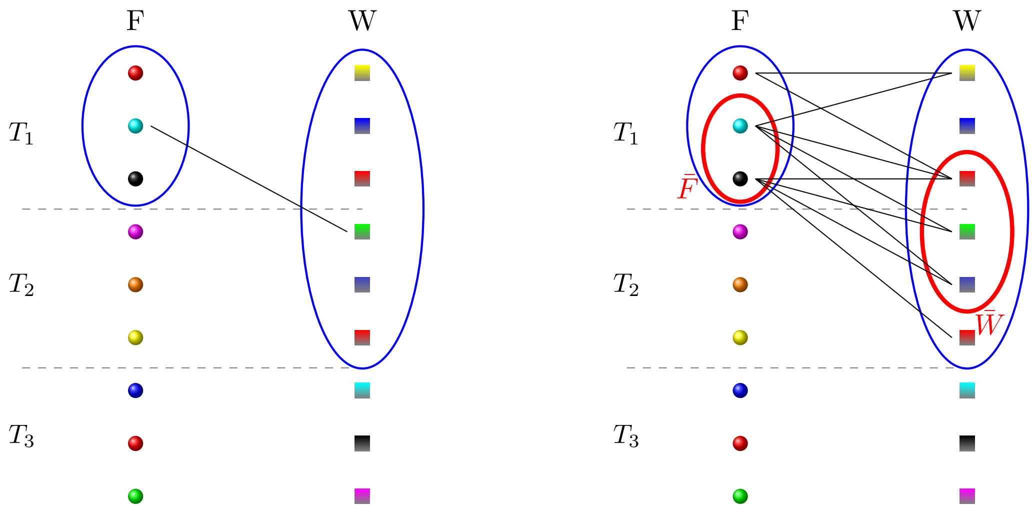


Figure 2: For each realized utility, we draw a bipartite graph with firms in tier-1 and workers in tiers up to 2 as the partitioned set of nodes (left). Firms in tier-1 receiving low utilities ($\bar{F}$) and workers in tiers up to 2 matched with non tier-1 firms ($\bar{W}$) form a biclique (right).

That is, f and w mutually fail to achieve high levels of utility.

On the other hand, since they are not joined by an edge,

$$\zeta_{f,w} > \bar{u} - \epsilon \quad \text{and} \quad \eta_{f,w} > \bar{v} - (v_1^o - v_2^o),$$

and therefore

$$u_{f,w} > u_2^o + \bar{u} - \epsilon \quad \text{and} \quad v_{f,w} > v_1^o + \bar{v} - (v_1^o - v_2^o) = v_2^o + \bar{v}.$$

In other words, the firm-worker pair's private-values are mutually so high that they would have achieved high utilities by making a blocking pair. This contradicts that μ_W is a stable matching.

This construction of a bipartite graph fits into a random bipartite graph model. Given that the tier-structure specifies a bi-partitioned set of nodes, we draw a bipartite graph based on the realized private-values. Since the private-values are i.i.d, each firm-worker pair is joined by an edge independently and with an identical probability. By Theorem 3, if the bi-partitioned set of nodes has a size on the order of n, and each pair of nodes is joined by an edge independently with a fixed probability, then the maximum balanced biclique has a size on the order of $\log(n)$ with a sequence of probabilities converging to 1 as n gets large. In addition, $\bar{W}$ contains at least $n/3$ workers, since there are $2n/3$ workers in tiers up to 2, but only $n/3$ firms in tier-1: i.e. $\bar{W}$ has a size on the order of n. Therefore, $\bar{F}$ must have a size, at most, on the order of $\log(n)$ with a sequence of probabilities converging to 1. The biclique $\bar{F} \cup \bar{W}$ would otherwise contain a balanced biclique with a size bigger than on the order of $\log(n)$, violating the Theorem 3. Lastly, $E\left[\frac{|\bar{F}|}{n/3}\right] \to 0$ follows immediately from $\log(n)/n \to 0$.

For the main theorem (without tier structure), we begin the proof by partitioning the supports of distributions for common-values. Suppose the common-values are drawn from the uniform distribution over $[0, 1]$. We partition the unit interval into K subintervals with equal lengths. Workers and firms are, in turn, grouped into tiers where firms or workers in the same tier have common-values in the same subinterval. Basically, we continue the proof as if we have a model with a finite number K of tiers. The tiers, though, need to be handled with care. This time, because the common-values are random, the tier structure is random. Moreover, agents in adjacent tiers may have arbitrarily close common-values.

As we increase the number of partitions K, the asymptotic lower bound on the utilities of firms in tier-k becomes close to the maximal utility achievable by matching with a worker in tier-k. With a similar exercise, we find an asymptotic lower bound on utilities of workers in each tier. Then, workers in tiers significantly higher than k are most likely to match with firms in tiers higher than k. This assortative-like feature of stable matchings induces an asymptotic upper bound on utilities of tier-k firms. As we finely partition the supports of the distributions of common-values, the differences in the common-values of firms or workers in similar tiers become slightly distinct from each other. Therefore, the asymptotic upper bound on utilities of firms in tier-k also becomes close to the maximal utility achievable by matching with a worker in tier-k. That is, we can find an asymptotic lower bound and an asymptotic upper bound, which are arbitrarily close to each other.

5 Market with Incomplete Information

We have so far considered a market with complete information. Agents are assumed to be able to assess the exact gain by misrepresenting preferences. It is a strong assumption, especially when we consider large markets. More realistically, we may want to consider a market with incomplete information, where each agent is only partially informed about the preferences of other participants. Nevertheless, we have mainly focused on the case of complete information since we can extrapolate its findings to show that the incentive to misrepresent preferences vanishes under incomplete information.

In relaxing the complete information assumption, we may consider various information structures. Each agent may know only the probability distributions in addition to either (i) her own utilities; (ii) her own utilities and the common-values of the other side; (iii) her own utilities, the common-values of the other side, and her own common-value evaluated by the other side; or (iv) her own utilities and all agents' common-values. The following results in the context of incomplete information correspond to the main theorem and its direct corollary for the model with complete information. As before, we first state the theorem informally, and then restate it with formal expressions.

Theorem *Regardless of information structure and for every $\epsilon > 0$, the expected proportion of firms (and workers) whose expected differences in utilities generated by all stable matchings are less than ϵ, converges to one as the market becomes large.*

Corollary *For any positive cost of misrepresenting preferences, if other agents truthfully reveal their preferences, the expected proportion of agents who have no incentive to manipulate a stable* *matching mechanism converges to one as the market becomes large.*

The intuition behind the theorem is clear. An expectation is a convex combination of all realizations. The expected difference between utilities from firm-optimal and worker-optimal stable matchings under incomplete information is simply a convex combination of the differences between utilities from the two stable matchings in all realized market instances. The differences between utilities are most likely to be insignificant (Theorem 1). Therefore, the expected difference in utilities is most likely to be negligible as well. We relegate the detailed proofs to Appendix D.

There are two advantages of showing the result in the context of complete information first, and then deriving the same result in the context of incomplete information. First, the results are robust to the information structure. The intuition of showing the results with incomplete information by using convex combinations remains valid regardless of the details of the information structure. Secondly, we can stress that non-manipulability of stable matching mechanisms is a property of the two-sided matching market itself, rather than stemming from insufficient information to manipulate the mechanism. Even when an agent can obtain complete knowledge of a preference profile at a small cost, it is not worth incurring that cost since the gain from manipulation will be small.

Formal Statement Let Π_f denote what f knows about a preference profile, and let π_f denote its realization. Then, the various incomplete information structures are denoted by (i) $\Pi_f = \langle U_{f,w} \rangle_{w \in W}$; (ii) $\Pi_f = \langle U_{f,w}, U^o_w \rangle_{w \in W}$; (iii) $\Pi_f = \langle U_{f,w}, U^o_w \rangle_{w \in W} \cup \{V^o_f\}$; and (iv) $\Pi_f = \langle U_{f,w}, U^o_w \rangle_{w \in W} \cup \langle V^o_{f'} \rangle_{f' \in F}$. Given a market instance $\langle F, W, u, v \rangle$, we define $\Delta_E(f; u, v)$ as the expected difference between utilities from firm-optimal and worker-optimal stable matchings conditioned on π_f. That is,

$$\Delta_E(f; u, v) := E_{U,V}\left[u_{\mu_F}(f) - u_{\mu_W}(f) \mid \pi_f\right],$$

where the expectation is applied to firm-optimal and worker-optimal stable matchings. For every $\epsilon > 0$, we correspondingly have the set of firms, whose expected differences in utilities from all stable matchings are less than ϵ, which is denoted by

$$A^F_E(\epsilon; u, v) := \{f \in F \mid \Delta_E(f; u, v) < \epsilon\}.$$

The previous theorem is an informal statement of the following theorem. We have similar notations and a theorem for workers, which are omitted here.

Theorem 4. *For any given information structure and for every $\epsilon > 0$,*

$$E\left[\frac{\left|A^F_E(\epsilon; U, V)\right|}{n}\right] \to 1, \quad as \quad n \to \infty.$$

Equilibrium Analysis Unfortunately, we do not obtain an equilibrium corresponding to the ϵ-Nash equilibrium in the context of complete information by using convex combinations. The obstacle to obtaining an equilibrium is that truncations by some agents may *increase* the differences in utilities generated by stable matchings for other participants. When preferences are known to all participants, truncations can preserve a stable matching under true preferences as individually rational under the announced preferences. The following example shows that this condition is necessary for truncations by some agents to decrease the differences in utilities from stable matchings for other participants.

$$
\begin{array}{llllll}
\mathbf{f_1}\!: & \langle w_1\rangle \succ w_2 \succ w_3 & & \mathbf{w_1}\!: & f_2 \succ \langle f_1\rangle \succ f_3 \\
\mathbf{f_2}\!: & \langle w_2\rangle \succ w_3 \succ w_1 & & \mathbf{w_2}\!: & f_1 \succ \langle f_2\rangle \succ f_3 \\
\mathbf{f_3}\!: & w_1 \succ w_2 \succ \langle w_3\rangle & & \mathbf{w_3}\!: & f_1 \succ f_2 \succ \langle f_3\rangle
\end{array}
,
$$

$$
\begin{array}{llllll}
\mathbf{f_1}\!: & \langle w_1\rangle \succ [w_2] \succ w_3 & & \mathbf{w_1}\!: & [f_2] \succ \langle f_1\rangle \succ f_3 \\
\mathbf{f_2}\!: & \langle w_2\rangle \succ w_3 \succ [w_1] & & \mathbf{w_2}\!: & [f_1] \succ \langle f_2\rangle \succ f_3 \\
\mathbf{f_3}\!: & w_1 \succ w_2 & & \mathbf{w_3}\!: & f_1
\end{array}
$$

Table 4: True preferences (upper) and their truncations (lower).

Table 4 lists true preferences of firms and workers (upper tables) and their truncations (lower tables). In the example, there is a unique stable matching (marked by $\langle\cdot\rangle$) under the true preferences. When f_3 and w_3 truncate their preferences, however, there are two stable matchings (marked by $\langle\cdot\rangle$ and $[\cdot]$). If some agents announce that all stable matching partners are unacceptable, other agents may have larger differences in utilities from all stable matchings.

Given incomplete information of a preference profile, an agent may submit a truncation of her true preference list based on the expected utility gain by manipulation. She may then remain unmatched afterwords depending on the realized preference profile. In this case, truncations may expand differences in utilities from stable matchings of other participants. Although most agents initially have small differences in utilities from stable matchings, participants may want to misrepresent their preferences as a best response to other agents' truncations.

6 Conclusions

This paper demonstrates an asymptotic similarity of stable matchings as the number of participants becomes large. Our measure of similarity is based on utilities, by which ordinal preferences are determined. As the utilities are drawn from some underlying probability distributions, one can analyze the likely differences in utilities from all stable matchings. We show that the expected proportion of firms and workers who are close to being indifferent among all stable partners converges to one as the market becomes large.

The result also implies that the expected proportion of agents who have a significant incentive to manipulate the mechanism vanishes in large markets. This is because the gain from manipulation of a stable matching mechanism is bounded above by the difference between utilities from the firm-optimal and the worker-optimal stable matchings. In addition, we show that with high probability a large market has an ϵ-Nash equilibrium in which most agents reveal their true preferences. We prove our results using techniques from the theory of random bipartite graphs, which is a new approach in the matching literature.

This paper is one of many recent studies exploring how the popularly used matching mechanisms really work in practice. It is essential to have a better understanding of stable matching mechanisms as market design applications expand from the NRMP and the School Choice Programs to many other markets, including dental residencies, various medical specialty matching programs, and labor markets for law clerks. Of particular relevance here is the fact that market designers are hoping to investigate the desirability of a clearinghouse in the market for economics Ph.D.s (Coles, Cawley, Levine, Niederle, Roth, and Siegfried, 2010). As such, understanding stable matching mechanisms in real applications becomes not only a market designers' question in theory, but is of concrete interest for economists in general.

Appendix

First in Appendix A, we summarize definitions and related theorems of asymptotic statistics. We prove Theorem 1 for the case of $\lambda = 0$ in Appendix B, and for the case of $0 < \lambda < 1$ in Appendix C. The proof of Theorem 4 is given in Appendix D. Lastly in Appendix E, we provide additional simulation results of effects of limited acceptability assumption on the proportion of unmatched agents.

A Asymptotic Statistics (Serfling, 1980)

Let $X_1, X_2, \ldots$ and X be random variables on a probability space $(\Omega, \mathcal{A}, P)$. We say that X_n **converges in probability** to X if

$$\lim_{n\to\infty} P\left(|X_n - X| < \epsilon\right) = 1, \quad \text{every } \epsilon > 0.$$

This is written $X_n \xrightarrow{p} X$.

For $r > 0$, we say that X_n **converges in the** r^{th} **mean** (or in the L^r-norm) to X if

$$\lim_{n\to\infty} E\left(|X_n - X|^r\right) = 0.$$

This is written $X_n \xrightarrow{L^r} X$.

Theorem A.1. *If* $X_n \xrightarrow{L^r} X$*, then* $X_n \xrightarrow{p} X$.

Theorem A.2. *Suppose that* $X_n \xrightarrow{p} X$*,* $|X_n| \le |Y|$ *with probability 1 (for all* n*), and* $E(|Y|^r) < \infty$*. Then,* $X_n \xrightarrow{L^r} X$.

Remark 1. *In this paper, most random variables represent proportions, which are bounded above by 1 with probability 1. As such, convergence in probability and convergence in the* r^{th} *mean are equivalent.*

Theorem A.3. *Let* $\mathbf{X_1}, \mathbf{X_2}, \ldots,$ *and* $\mathbf{X}$ *be random* k*-vectors defined on a probability space, and let* g *be a vector-valued Borel function defined on* $\mathbf{R}^k$*. If* g *is continuous with* $P_{\mathbf{X}}$*-probability 1, then*

$$\mathbf{X_n} \xrightarrow{p} \mathbf{X} \implies g(\mathbf{X_n}) \xrightarrow{p} g(\mathbf{X}).$$

In particular, if $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} Y$, then $X_n + Y_n \xrightarrow{p} X + Y$ and $X_n Y_n \xrightarrow{p} XY$.

Given a univariate distribution function F and $0 < q < 1$, we define q^{th} **quantile** ξ_q as

$$\xi_q := \inf\{x : F(x) \ge q\}.$$

Consider an i.i.d sequence $\langle X_i \rangle$ with distribution function F. For each sample of size n, $\{X_1, X_2, \ldots, X_n\}$, a corresponding **empirical distribution function** F_n is constructed as

$$F_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i \le x\}, \quad -\infty < x < \infty.$$

The **empirical** q^{th} **quantile** $\hat{\xi}_{q:n}$ is defined as the q^{th} quantile of the empirical distribution function. That is

$$\hat{\xi}_{q:n} := \inf\{x : F_n(x) \ge q\}.$$

For each x, $F_n(x)$ is a random variable, and therefore, $\hat{\xi}_{q:n}$ is also a random variable.

Theorem A.4. *Suppose that* q^{th} *quantile* ξ_q *is the unique solution* x *of* $F(x-) \le q \le F(x)$*. Then, for every* $0 < q < 1$ *and* $\epsilon > 0$*,*

$$P\left(\left|\hat{\xi}_{q:n} - \xi_q\right| > \epsilon\right) \le 2e^{-2n\lambda_\epsilon^2}$$

for all n*, where* $\lambda_{1,\epsilon} = F(\xi_q + \epsilon) - q$*,* $\lambda_{2,\epsilon} = q - F(\xi_q - \epsilon)$*, and* $\lambda_\epsilon = \min\{\lambda_{1,\epsilon},\ \lambda_{2,\epsilon}\}$.

For each sample of size n, $\{X_1, X_2, \ldots, X_n\}$, the ordered sample values

$$X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$$

are called the **order statistics**.

In view of

$$X_{k:n} = \hat{\xi}_{k/n:n}, \quad 1 \le k \le n, \tag{1}$$

we will carry out proofs in terms of empirical quantiles, even when variables are defined as order statistics.

B Proof of Theorem 1 ($\lambda = 0$)

Let $\zeta = [\zeta_{f,w}]$ be an i.i.d sample from a continuous distribution Γ^W with support $[0, \bar{u}]$, and $\eta = [\eta_{f,w}]$ be an i.i.d sample from a continuous distribution Γ^F with support $[0, \bar{v}]$.[19]

For $\epsilon > 0$ and for each $\langle F, W, u, v \rangle$, we define

$$B^F(\epsilon; u, v) := F \backslash A^F(\epsilon; u, v) = \{f \in F \mid \Delta(f; u, v) \geq \epsilon\},$$

and prove that

$$E\left[\frac{|B^F(\epsilon; U, V)|}{n}\right] \to 0, \quad \text{as} \quad n \to \infty. \tag{2}$$

We define the set of firms whose utilities from the worst stable matching are significantly below the upper bound $\bar{u}$, which we shall write as

$$\bar{B}(\epsilon; u, v) := \{f \in F | u_{\mu_W}(f) \leq \bar{u} - \epsilon\}.$$

Note from $u_{\mu_F}(f) \leq \bar{u}$ that

$$u_{\mu_F}(f) - u_{\mu_W}(f) \leq \bar{u} - u_{\mu_W}(f),$$

and thus

$$B^F(\epsilon; u, v) \subset \bar{B}(\epsilon; u, v).$$

Therefore, (2) follows immediately from the following proposition.

Proposition B.1. *For every $\epsilon > 0$,*

$$E\left[\frac{|\bar{B}(\epsilon; U, V)|}{n}\right] \to 0 \quad \text{as} \quad n \to \infty.$$

We divide the proof of Proposition B.1 into two lemmas. For every market instance $\langle F, W, u, v \rangle$, we let $R_{\mu_W}(f)$ be the rank number of firm f's worker-optimal stable matching partner: e.g. $R_{\mu_W}(f) = 1$ if f matches with its most preferred worker. We first observe that for most firms, the rank number of worker-optimal matching partner normalized by n converges to 0. The second lemma shows that the corresponding utility level must become close to the upper bound $\bar{u}$ as the market becomes large.

[19] We use Γ^W, instead of Γ^F, to represent the distribution of utilities of firms, interpreting it as the distribution of private-values of workers. This notation will be consistent with the additional notation G^W representing the distribution of workers' common-values. By the same reason, we use Γ^F to denote the distribution of utilities of workers, or private-values of firms.

Lemma B.2. *For $\gamma > 0$, let*

$$\bar{B}_q(\gamma; u, v) := \left\{ f \in F \mid \frac{R_{\mu_W}(f)}{n} \geq \gamma \right\} = \left\{ f \in F \mid 1 - \frac{R_{\mu_W}(f)}{n} \leq 1 - \gamma \right\}.$$

Then, for every sequence $\langle \gamma_n \rangle$ such that $\gamma_n \to 0$ and $(\log n) \cdot \gamma_n \to \infty$,

$$E\left[\frac{\left|\bar{B}_q(\gamma_n; U, V)\right|}{n} \right] \to 0 \quad as \quad n \to \infty.$$

Proof. For every instance $\langle F, W, u, v \rangle$ and for every sequence $\langle \gamma_n \rangle$ satisfying the conditions,

$$\begin{aligned} \frac{1}{n}\gamma_n \left|\bar{B}_q(\gamma_n; u, v)\right| &\leq \frac{1}{n} \sum_{f \in \bar{B}_q(\gamma_n; u, v)} \frac{R_{\mu_W}(f)}{n} \\ &\leq \frac{1}{n} \sum_{f \in F_n} \frac{R_{\mu_W}(f)}{n}. \end{aligned}$$

We use Theorem 2 in Pittel (1989) showing that

$$\frac{\sum_{f \in F_n} R_{\mu_W}(f)}{n^2 \log^{-1} n} \xrightarrow{p} 1. \tag{3}$$

Applying (3), we shall write

$$\begin{aligned} \frac{\left|\bar{B}_q(\gamma_n; U, V)\right|}{n} &\leq \frac{\sum_{f \in F_n} R_{\mu_W}(f)}{n^2} \frac{1}{\gamma_n} \\ &= \frac{\sum_{f \in F_n} R_{\mu_W}(f)}{n^2 \log^{-1} n} \frac{1}{\log n \cdot \gamma_n} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty. \end{aligned}$$

We obtain Lemma B.2 since $\frac{\left|\bar{B}_q(\gamma_n; U, V)\right|}{n}$ is bounded above by 1 with probability 1 for all n so that convergence in probability implies convergence in mean (Theorem A.2). □

Lemma B.3. *For any $\gamma > 0$, let*

$$\bar{B}'(\epsilon, 1 - \gamma; u, v) := \left\{ f \in F \mid \hat{\xi}^f_{1-\gamma; n} \leq \bar{u} - \epsilon \right\},$$

where $\hat{\xi}^f_{1-\gamma; n}$ is the realized value of the empirical $(1 - \gamma)^{th}$ quantile of $U_f = \langle U_{f,w} \rangle_{w \in W_n}$.

Then, for every $\epsilon > 0$ and sequence $\langle \gamma_n \rangle$ such that $\gamma_n \to 0$ and $(\log n) \cdot \gamma_n \to \infty$,

$$E\left[\frac{\left|\bar{B}'(\epsilon, 1 - \gamma_n; U, V)\right|}{n} \right] \to 0 \quad as \quad n \to \infty.$$

Proof. For each n, let $f_n \in F_n$ and consider the resulting sequence $\langle f_n \rangle_{n=1}^{\infty}$. Note that

$$E\left[\frac{|\bar{B}'(\epsilon, 1-\gamma_n; U, V)|}{n}\right] = \frac{1}{n} \sum_{f \in F_n} E\left[\mathbf{1}\left\{\hat{\xi}^{f}_{1-\gamma_n;n} \leq \bar{u} - \epsilon\right\}\right]$$
$$= E\left[\mathbf{1}\left\{\hat{\xi}^{f_n}_{1-\gamma_n;n} \leq \bar{u} - \epsilon\right\}\right]$$
$$= P\left(\hat{\xi}^{f_n}_{1-\gamma_n;n} \leq \bar{u} - \epsilon\right).$$

Thus, it is enough to show that

$$P\left(\hat{\xi}^{f_n}_{1-\gamma_n;n} \leq \bar{u} - \epsilon\right) \to 0, \quad \text{as} \quad n \to \infty.$$

Take any q from the interval $\big(\Gamma^W(\bar{u}-\epsilon), 1\big)$ such that q^{th} quantile ξ_q is the unique solution x of $\Gamma^W(x-) \leq q \leq \Gamma^W(x)$.[20] For any large n, we have $1 - \gamma_n > q$, and thus

$$\hat{\xi}^{f_n}_{1-\gamma_n;n} \geq \hat{\xi}^{f_n}_{q;n}.$$

Therefore, we shall write

$$P\left(\hat{\xi}^{f_n}_{1-\gamma_n;n} \leq \bar{u} - \epsilon\right) \leq P\left(\hat{\xi}^{f_n}_{q;n} \leq \bar{u} - \epsilon\right)$$
$$= P\left(\left|\hat{\xi}^{f_n}_{q;n} - \xi_q\right| \geq \xi_q - (\bar{u} - \epsilon)\right),$$

which converges to 0 by Theorem A.4. □

We complete the proof of Proposition B.1 using the following observation. For each $\langle F, W, u, v \rangle$ and for every sequence $\langle \gamma_n \rangle$ such that $\gamma_n \to 0$ and $(\log n) \cdot \gamma_n \to \infty$,

$$\bar{B}(\epsilon; u, v) = \big(\bar{B}(\epsilon; u, v) \cap \bar{B}_q(\gamma_n; u, v)\big) \cup \big(\bar{B}(\epsilon; u, v) \cap (F \backslash \bar{B}_q(\gamma_n; u, v))\big)$$
$$\subset \bar{B}_q(\gamma_n; u, v) \cup \big(\bar{B}(\epsilon; u, v) \cap (F \backslash \bar{B}_q(\gamma_n; u, v)\big).$$

Each f in $F \backslash \bar{B}_q(\gamma_n; u, v)$ matches in μ_W with a worker of a normalized rank less than γ_n. Nevertheless if f obtains utility less than $\bar{u} - \epsilon$ in μ_W (i.e. $f \in \bar{B}(\epsilon; u, v)$), then the realized empirical $(1-\gamma_n)^{th}$ quantile of his utilities is below $\bar{u} - \epsilon$.

That is,

$$\bar{B}(\epsilon; u, v) \cap F \backslash \bar{B}_q(\gamma_n; u, v) \subset \bar{B}'(\epsilon, 1 - \gamma_n; u, v),$$

[20] There exists such a q. For every q in $\big(\Gamma^W(\bar{u}-\epsilon), 1\big)$, we have x_q in $(\bar{u}-\epsilon, \bar{u})$ such that $\Gamma^W(x_q) = q$ by Intermediate Value Theorem. Suppose toward contradiction that every q has two distinct $\underline{x}_q$ and $\bar{x}_q$ in $(\bar{u}-\epsilon, \bar{u})$ such that $\Gamma^W(\underline{x}_q) = \Gamma^W(\bar{x}_q) = q$. Since $\Gamma^W(\cdot)$ is a distribution, every q then has a closed interval $[\underline{x}_q, \bar{x}_q]$ such that $\Gamma^W(x) = q$ for all $x \in [\underline{x}_q, \bar{x}_q]$. Moreover, if $q \neq q'$, then $[\underline{x}_q, \bar{x}_q]$ and $[\underline{x}_{q'}, \bar{x}_{q'}]$ are disjoint. There are uncountable number of elements in $\big(\Gamma^W(\bar{u}-\epsilon), 1\big)$, whereas there are at most countable number of closed disjoint intervals in $(\bar{u}-\epsilon, \bar{u})$.

and therefore

$$\bar{B}(\epsilon; u, v) \subset \bar{B}_q(\gamma_n; u, v) \cup \bar{B}'(\epsilon, 1-\gamma_n; u, v).$$

We proved in Lemma B.2 and B.3 that both $\frac{|\bar{B}_q(\gamma_n;U,V)|}{n}$ and $\frac{|\bar{B}'(\epsilon,1-\gamma_n;U,V)|}{n}$ converge to 0 in mean, which completes the proof.

C Proof of Theorem 1 ($0 < \lambda < 1$).

To simplify notations, we compress λ and $1-\lambda$, and consider utilities defined as

$$U_{f,w} = U^o_w + \zeta_{f,w} \quad \text{and} \quad V_{f,w} = V^o_f + \eta_{f,w}.$$

We do not lose generality since we can regard common-values and private-values as the ones already multiplied by λ and $1-\lambda$, respectively.

Let U^o_n and V^o_n be i.i.d samples of size n from distributions G^W and G^F, respectively. G^W and G^F have strictly positive density functions on supports in $\mathbb{R}_+$.[21] $\zeta = [\zeta_{f,w}]$ is an i.i.d sample from a continuous distribution Γ^W with support $[0, \bar{u}]$, and $\eta = [\eta_{f,w}]$ is an i.i.d sample from a continuous distribution Γ^F with support $[0, \bar{v}]$.

We define

$$B^F(\epsilon; u, v) := F \backslash A^F(\epsilon; u, v) = \{f \in F \mid \Delta(f; u, v) \geq \epsilon\}$$

and prove that $\frac{|B^F(\epsilon;U,V)|}{n}$ converges to 0 in probability, which is equivalent to proving convergence to 0 in mean (Theorem A.2). That is, we fix $\epsilon > 0$ and $K \in \mathbb{N}$, and prove that

$$P\left(\frac{|B^F(\epsilon; U, V)|}{n} > \frac{9}{K}\right) \to 0, \quad \text{as} \quad n \to \infty.$$

First, we partition the supports of the common-value distributions into K intervals. Then for each market instance, in particular for each realized profile of common-values, we group firms and workers into *two* versions of a finite number of tiers, where agents in the same tier have similar common-values. We first find that tier-k firms are most likely to achieve a utility level higher than an arbitrary ϵ less than the maximal utility achievable from workers in tier-$(k+3)$ (Proposition C.1).[22] For the proof, we use techniques from a theory of random bipartite graphs.

Once we find an asymptotic lower bound on utilities of firms in each tier, we find an asymptotic upper bound on utilities of firms in a tier, say k, by referencing to the asymptotic lower bounds

[21] The bounded support condition of common-value distributions is not necessary for Theorem 1. We use the condition later for Theorem 4 of incomplete information case.

[22] In Section 4, we showed with a market with tiers that firms in tier-t are most likely to achieve a utility level higher than an arbitrary ϵ less than the maximal utility from a worker in tier-$(k+1)$. In the model with tiers, each tier has a distinct tier-specific common-value, so there is a clear-cut distinction between tier-k and tier-$(k+1)$ specific values. In the general model (without tiers), however, there is no such distinction in common-values between adjacent tiers. The highest common-value of workers in tier-$(k+1)$ can be arbitrarily close to the lowest common-value of workers in tier-k. This leads us to use the maximal utility from a worker in tier-$(k+3)$ rather than tier-$(k+1)$ as an asymptotic lower bound on utilities of tier-k firms.

on utilities of workers in tiers higher than k (Proposition C.2). As workers in higher tiers achieve high utilities, they are most likely to match with firms in high tiers, rather than firms in tier-k. Accordingly, the utilities of tier-k firms are asymptotically bounded above by the maximal utility that they can achieve by matching with workers in tiers near k.

As we finely partition the supports of the common-value distributions, the differences in common-values between adjacent tiers become small. Then, the asymptotic lower bound on utilities of tier-k firms will become close to the maximal utility achievable from workers in tier-k. In addition, the asymptotic upper bound also becomes close to the same level, since the maximal utility achievable from workers in tiers near k will also become close to the maximal utility achievable from workers in tier-k.

We divide the proof into three subsections. First, in Subsection C.1, we construct *two* tier structures from each profile of realized common-values. Then in Subsection C.2, we define three events related to the tier structures, and show that the all three events occur with a probability converging to 1 as the market becomes large. The real proof begins in subsection C.3. During the proof, we shall focus on the market instances where realized firms' or workers' common-values are all distinct. G^F and G^W are continuous, ensuring that realized common-values are all distinct with probability 1.

C.1 Tier-Grouping

We use the following notations.

1. ξ_q^F and ξ_q^W : q^{th} quantile of G^F and G^W.
2. $\hat{\xi}_{q;n}^F$ and $\hat{\xi}_{q;n}^W$: empirical q^{th} quantile of n-size samples from distributions G^F and G^W, respectively. We also use $\hat{\xi}_{q;n}^F$ and $\hat{\xi}_{q;n}^W$ to denote their realizations.

Since realized common-values $u_n^o = \langle u_w^o \rangle_{w \in W_n}$ and $v_n^o = \langle v_f^o \rangle_{f \in F_n}$ are all distinct with probability 1, we index firms and workers from $i = 1$ to n in the order of their common-values: i.e.

$$v_{f_i}^o > v_{f_j}^o \quad \text{and} \quad u_{w_i}^o > u_{w_j}^o, \quad \text{if} \quad i < j.$$

Then, $U_{w_i;n}^o$ and $V_{f_i;n}^o$ represent i^{th} highest values of n order statistics from G^W and G^F. Note that $U_{w_i;n}^o = \hat{\xi}_{(1-\frac{i-1}{n});n}^W$ by the relationship between order statistics and empirical quantiles (see Equation (1)).

We partition the support of G^W into

$$\begin{aligned}
I_1^W &:= (\xi^W_{1-\frac{1}{K}}, \infty] \\
I_2^W &:= (\xi^W_{1-\frac{2}{K}}, \xi^W_{1-\frac{1}{K}}] \\
&\vdots \\
I_k^W &:= (\xi^W_{1-\frac{k}{T}}, \xi^W_{1-\frac{k-1}{K}}] \\
&\vdots \\
I_K^W &:= [0, \xi^W_{\frac{1}{K}}].
\end{aligned}$$

We define **the set of workers in tier-k** (with respect to *workers'* common-values) as

$$W_k(u) := \left\{ w \mid u^o_w \in I^W_k \right\} \quad \text{for} \quad k = 1, 2, \ldots, K,$$

and define **the set of firms in tier-k** (with respect to *workers'* common-values) as

$$F_k(u) := \{f_i \in F_n \mid w_i \in W_k(u)\}.$$

We will use the following notations.

1. $l_k(u) := |F_k(u)| = |W_k(u)|$: The size of tier-$k$ (with respect to *workers'* common-values).
2. $u^o_k := \xi^W_{1-\frac{k}{K}}$: The threshold level of tier-k and tier-$(k+1)$ workers' common-values. Note, $w \in W_k(u)$ if and only if $u^o_k < u^o_w \leq u^o_{k-1}$.

Remark 2. *The set of tier-k workers is defined with respect to workers' common-values, which is a random sample. Therefore, $W_k(U)$ is random, and so is $F_k(U)$. In particular, the size of tier-k, $l_k(U)$, is random; whereas, u^o_k is a constant.*

In parallel, we partition the support of G^F into

$$\begin{aligned}
I_1^F &:= (\xi^F_{1-\frac{1}{K}}, \infty] \\
I_2^F &:= (\xi^F_{1-\frac{2}{K}}, \xi^F_{1-\frac{1}{K}}] \\
&\vdots \\
I_k^F &:= (\xi^F_{1-\frac{k}{K}}, \xi^F_{1-\frac{k-1}{K}}] \\
&\vdots \\
I_K^F &:= [0, \xi^F_{\frac{1}{K}}].
\end{aligned}$$

We define **the set of firms in tier-k** (with respect to *firms'* common-values) as

$$F_k(v) := \left\{ f \mid v^o_f \in I^F_k \right\} \quad \text{for} \quad k = 1, 2, \ldots, K,$$

and define **the set of workers in tier-k** (with respect to *firms'* common-values) as

$$W_k(v) := \{w_i \in W_n \mid f_i \in F_k(v)\}.$$

Accordingly, we use the following notations.

1. $l_k(v) := |F_k(v)| = |W_k(v)|$: The size of tier-$k$ (with respect to *firms'* common-values).
2. $v_k^o := \xi^F_{1-\frac{k}{K}}$: The threshold level of tier-k and tier-$(k+1)$ firms' common-values. Note, $f \in F_k(u)$ if and only if $v_k^o < v_f^o \leq v_{k-1}^o$.

Remark 3. *Tiers with respect to workers' common-values are in general not the same as tiers with respect to firms' common-values. In particular, we are most likely to have $l_k(u) \neq l_k(v)$.*

Throughout the proof, we mainly use tiers defined with respect to workers' common-values. However, we need both tier structures in the last part of the proof. We simply write "tier-k" to denote tier-k with respect to workers' common-values, and use "(w.r.t firm) tier-k" to denote tier-k with respect to firms' common-values.

C.2 High-Probability Events

We introduce three events and show that the events occur with probabilities converging to 1 as the market becomes large. We provide proofs for completeness, but the main ideas are simply from the (weak) law of large numbers. In the next section, we will leave the probability that the following events do not occur as a remainder term converging to zero, and focus on the probabilities conditioned that the following events all occur.

C.2.1 No vanishing tiers

Event 1 ($\mathcal{E}_1$)**.** *Let $\bar{K} > K$. For all $k = 1, 2, \ldots, K$,*

$$\frac{l_k(U)}{n} > \frac{1}{\bar{K}}.$$

Proof. By definition,

$$\frac{l_k(U)}{n} := \frac{1}{n} \sum_{w \in W_n} \mathbf{1}\{U_w^o \in I_k^W\},$$

which converges to $\frac{1}{K}$ in probability by the (weak) law of large numbers. □

C.2.2 Distinct common-values of the firms in non-adjacent tiers.

Let $\tilde{\epsilon} > 0$ be such that for any $v, v' \in [0, \xi^F_{1-1/K}]$ and $|v - v'| \leq \tilde{\epsilon}$,

$$|G^F(v) - G^F(v')| < \frac{1}{3K}.$$

There exists such an $\tilde{\epsilon}$ since G^F is uniformly continuous on $[0, \xi^F_{1-1/K}]$.

Event 2 ($\mathcal{E}_2$)**.** *For every* $k = 1, 2, \ldots, K-2$,

$$\min_{\substack{f \in F_k(U) \\ f' \in F_{k+2}(U)}} |V^o_f - V^o_{f'}| > \tilde{\epsilon}.$$

Proof. Fix $k \in 1, 2, \ldots, K-2$ and realized u. For every $w_i \in W_k(u)$ and $w_j \in W_{k+2}(u)$,

$$u^o_{w_i} > u^o_k = \xi^W_{1-\frac{k}{K}}, \quad \text{and} \quad u^o_{w_j} \leq u^o_{k+1} = \xi^W_{1-\frac{k+1}{K}}. \tag{4}$$

For any $q \in (0,1)$, $\hat{\xi}^W_{q;n} \xrightarrow{p} \xi^W_q$ (Theorem A.4), from which the following inequalities hold with probability converging to 1 as $n \to \infty$.

$$\xi^W_{1-\frac{k}{K}} > \hat{\xi}^W_{1-\frac{k}{K}-\frac{1}{4K}} \quad \text{and} \quad \xi^W_{1-\frac{k+1}{K}} < \hat{\xi}^W_{1-\frac{k+1}{K}+\frac{1}{4K}}. \tag{5}$$

Considering (4) and the relation between order statistics and empirical quantiles (Equation (1)), if (5) holds, we have

$$1 - \frac{k}{K} - \frac{1}{4K} < \min_{w_i \in W_k(u)} \left(1 - \frac{i-1}{n}\right) = \min_{f_i \in F_k(u)} \left(1 - \frac{i-1}{n}\right)$$

and

$$1 - \frac{k+1}{K} + \frac{1}{4K} > \max_{w_j \in W_{k+2}(u)} \left(1 - \frac{j-1}{n}\right) = \max_{f_j \in F_{k+2}(u)} \left(1 - \frac{j-1}{n}\right).$$

Then for every $f_i \in F_k(u)$ and $f_j \in F_{k+2}(u)$,

$$v^o_{f_i} > \hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{4K}} \quad \text{and} \quad v^o_{f_j} < \hat{\xi}^F_{1-\frac{k+1}{K}+\frac{1}{4K}}.$$

Therefore,

$$\begin{aligned} P\Big(\inf_{\substack{f_i \in F_k(U) \\ f_j \in F_{k+2}(U)}} \left|V^o_{f_i} - V^o_{f_j}\right| \leq \tilde{\epsilon}\Big) &\leq P\left(\left|\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{4K}} - \hat{\xi}^F_{1-\frac{k+1}{K}+\frac{1}{4K}}\right| \leq \tilde{\epsilon}\right) + R_n \\ &\leq P\left(\left|G^F(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{4K}}) - G^F(\hat{\xi}^F_{1-\frac{k+1}{K}+\frac{1}{4K}})\right| < \frac{1}{3K}\right) + R_n, \end{aligned} \tag{6}$$

where R_n corresponds to the probability that (5) is violated: i.e. $R_n \to 0$. The last inequality is by the definition of $\tilde{\epsilon}$.

Note that

$$G^F(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{4K}}) - G^F(\hat{\xi}^F_{1-\frac{k+1}{K}+\frac{1}{4K}}) \xrightarrow{p} \frac{1}{2K}$$

by Theorem A.4 and continuity of G^F(Theorem A.3). As a result, the right hand side of (6) converges to 0.

□

C.2.3 Similarity between tiers w.r.t workers' common-values and tiers w.r.t firms' common-values

The following event is the case that all firms in tier-k with respect to workers' common-values are in a tier near k with respect to firms' common-values, and vice versa.

Event 3 ($\mathcal{E}_3$)**.** *For every $k = 1, 2, 3, \ldots, K$,*

$$F_k(U) \subset \bigcup_{k'=k-1}^{k+1} F_{k'}(V) \quad \text{and} \quad W_k(V) \subset \bigcup_{k'=k-1}^{k+1} W_{k'}(U).^{23}$$

Proof. We prove the first part and omit the proof of the second part.

For each realized (u, v), we have

$$\{u^o_w | w \in W_k(u)\} \subset \left(u^o_k, u^o_{k-1}\right] = \left(\xi^W_{1-\frac{k}{K}}, \xi^W_{1-\frac{k-1}{K}}\right]. \tag{7}$$

Suppose

$$\left(\xi^W_{1-\frac{k}{K}}, \xi^W_{1-\frac{k-1}{K}}\right] \subset \left(\hat{\xi}^W_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^W_{1-\frac{k-1}{K}+\frac{1}{2K}}\right], \tag{8}$$

and

$$\left(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^F_{1-\frac{k-1}{K}+\frac{1}{2K}}\right] \subset \left(\xi^F_{1-\frac{k+1}{K}}, \xi^F_{1-\frac{k-2}{K}}\right]. \tag{9}$$

If (8) hold, then (7) implies that for every tier-k worker w_i, we have

$$u^o_{w_i} \in \left(\hat{\xi}^W_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^W_{1-\frac{k-1}{K}+\frac{1}{2K}}\right],$$

and thus,

$$1 - \frac{i-1}{n} \in \left(1 - \frac{k}{K} - \frac{1}{2K}, 1 - \frac{k-1}{K} + \frac{1}{2K}\right].$$

Then for any tier-k firm f_i, we have

$$v^o_{f_i} \in \left(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^F_{1-\frac{k-1}{K}+\frac{1}{2K}}\right],$$

which implies that

$$\left\{v^o_f \mid f \in F_k(u)\right\} \subset \left(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^F_{1-\frac{k-1}{K}+\frac{1}{2K}}\right].$$

Consequently if both (8) and (9) hold, then

$$\begin{aligned}\left\{v^o_f \mid f \in F_k(u)\right\} &\subset \left(\hat{\xi}^F_{1-\frac{k}{K}-\frac{1}{2K}}, \hat{\xi}^F_{1-\frac{k-1}{K}+\frac{1}{2K}}\right] \\ &\subset \left(\xi^F_{1-\frac{k+1}{K}}, \xi^F_{1-\frac{k-2}{K}}\right] \\ &= \bigcup_{k'=k-1}^{k+1} I^F_{k'}.\end{aligned}$$

[23] We simply assume that $F_0(V)$, $F_0(V)$, $W_{K+1}(U)$, and $W_{K+1}(U)$ are empty sets.

In other words,

$$F_k(u) \subset \bigcup_{k'=k-1}^{k+1} F_{k'}(v).$$

(8) and (9) occur with probability converging to 1 (Theorem A.4), and thus the event $\mathcal{E}_3$ also occurs with probability converging to 1. $\square$

C.3 Proof of the Theorem 1

We choose K large enough that

$$\max_{1\le k\le K-1} \left|u_k^o - u_{k+1}^o\right| \equiv \max_{1\le k\le K-1} \left|\xi^W_{1-\frac{k}{K}} - \xi^W_{1-\frac{k+1}{K}}\right| < \frac{\epsilon}{9}.^{24} \tag{10}$$

We divide the proof into two propositions. The first proposition finds an asymptotic lower bound on utilities of firms in each tier, using techniques from the theory of random bipartite graphs. Similarly, we have a proposition for an asymptotic lower bound on utilities of workers in each tier. The second proposition derives an asymptotic upper bound on utilities of firms in each tier, by referencing the lower bounds on utilities of workers in higher tiers. The Theorem 1 follows from the fact that the lower bound and the upper bound are close to each other.

Proposition C.1. *For each instance $\langle F, W, u, v\rangle$ and for each $\bar{k} = 1, 2, \ldots, K-2$, define*

$$\hat{B}^F_{\bar{k}}(\epsilon; u, v) := \left\{ f \in F_{\bar{k}}(u) : u_{\mu_W}(f) \le u^o_{\bar{k}+2} + \bar{u} - \epsilon \right\}.^{25}$$

Then for any $\epsilon > 0$,

$$\frac{|\hat{B}^F_{\bar{k}}(\epsilon; U, V)|}{n} \xrightarrow{p} 0 \quad as \quad n \to \infty.$$

Proof. For each instance $\langle F, W, u, v\rangle$ and for each $k = 1, 2, \ldots, K$, let $F_{\le k}(u) := \bigcup_{k'\le k} F_{k'}(u)$ and $F_{<k}(u) := \bigcup_{k'<k} F_{k'}(u)$. Similarly, we define $W_{\le k}(u)$ and $W_{<k}(u)$.

Take any $\bar{k}$ from $\{1, 2, \ldots, K-2\}$. We construct a bipartite graph with $F_{\bar{k}}(u) \cup W_{\le \bar{k}+2}(u)$ as a partitioned set of nodes. (see Section 3 for the related definitions.) Two vertices $f \in F_{\bar{k}}(u)$ and $w \in W_{\le \bar{k}+2}(u)$ are joined by an edge if and only if

$$\zeta_{f,w} \le \bar{u} - \epsilon \quad \text{or} \quad \eta_{f,w} \le \bar{v} - \tilde{\epsilon},$$

where $\tilde{\epsilon}$ is the value taken before, while defining $\mathcal{E}_2$.

Let $\bar{W}_{\le \bar{k}+2}(u, v)$ be the set of workers in tiers up to $\bar{k}+2$ who are not matched with firms in tiers up to $\bar{k}+1$ in μ_W. That is,

$$\bar{W}_{\le \bar{k}+2}(u, v) := \left\{ w \in W_{\le \bar{k}+2}(u) \,|\, \mu_W(w) \notin F_{\le \bar{k}+1}(u) \right\}.$$

[24] We can always satisfy the condition since G^W has a strictly positive density function.

[25] Note that $u^o_{\bar{k}+2} + \bar{u}$ is the maximal utility level a firm can achieve by matching with a worker in tier-($\bar{k}+3$).

We now show that if $\mathcal{E}_2$ holds, then

$$\hat{B}^F_{\bar{k}}(\epsilon; u, v) \cup \bar{W}_{\leq \bar{k}+2}(u, v)$$

is a biclique.

Suppose, towards a contradiction, that a pair of $f \in \hat{B}^F_{\bar{k}}(\epsilon; u, v)$ and $w \in \bar{W}_{\leq \bar{k}+2}(u, v)$ is *not* joined by an edge: i.e.

$$\zeta_{f,w} > \bar{u} - \epsilon \quad \text{and} \quad \eta_{f,w} > \bar{v} - \tilde{\epsilon}.$$

Then, we first have

$$u_{f,w} = u^o_w + \zeta_{f,w} > u^o_{\bar{k}+2} + \zeta_{f,w} > u^o_{\bar{k}+2} + \bar{u} - \epsilon, \tag{11}$$

and also have

$$v_{f,w} = v^o_f + \eta_{f,w} \geq \min_{f' \in F_{\bar{k}}(u)} v^o_{f'} + \eta_{f,w} > \min_{f' \in F_{\bar{k}}(u)} v^o_{f'} + \bar{v} - \tilde{\epsilon}.$$[26]

Conditioned on $\mathcal{E}_2$, we can proceed further and obtain

$$\begin{aligned} v_{f,w} &> \min_{f' \in F_{\bar{k}}(u)} v^o_{f'} + \bar{v} - \left(\min_{f' \in F_{\bar{k}}(u)} v^o_{f'} - \max_{f'' \in F_{\bar{k}+2}(u)} v^o_{f''} \right) \\ &= \max_{f'' \in F_{\bar{k}+2}(u)} v^o_{f''} + \bar{v}. \end{aligned} \tag{12}$$

On the other hand, $f \in \hat{B}^F_{\bar{k}}(\epsilon; u, v)$ implies that

$$u_{\mu_W}(f) \leq u^o_{\bar{k}+2} + \bar{u} - \epsilon,$$

and $w \in \bar{W}_{\leq \bar{k}+2}(u, v)$ implies that

$$v_{\mu_W}(w) \leq \max_{f'' \in F_{\bar{k}+2}(u)} v^o_{f''} + \bar{v},$$

since a worker can obtain utility higher than $\max_{f'' \in F_{\bar{k}+2}(u)} v^o_{f''} + \bar{v}$ only by matching with a firm in $F_{\leq \bar{k}+1}(u)$.

Then, (11) and (12) implies that (f, w) would have blocked μ_W, contradicting that μ_W is stable. Therefore,

$$\hat{B}^F_{\bar{k}}(\epsilon; u, v) \cup \bar{W}_{\leq \bar{k}+2}(u, v).$$

is a biclique, which is not necessarily balanced.

We now control the size of $\hat{B}^F_{\bar{k}}(\epsilon; U, V)$ by referencing Theorem 3. Let u^o and v^o be realized common-values such that events $\mathcal{E}_1$ and $\mathcal{E}_2$ hold. Then, the remaining randomness of U and V is from ζ and η. Consider a random bipartite graph with $F_{\bar{k}}(U) \cup W_{\leq \bar{k}+2}(U)$ as a bi-partitioned set

[26] We should not replace $\min_{f' \in F_{\bar{k}}(u)} v^o_{f'}$ with $v^o_{\bar{k}}$. $F_{\bar{k}}(u)$ is defined with respect to workers' common-values, rather than firms' common-values.

of nodes, where each pair of $f \in F_{\bar{k}}(U)$ and $w \in W_{\leq \bar{k}+2}(U)$ is joined by an edge if and only if

$$\zeta_{f,w} \leq \bar{u} - \epsilon \quad \text{or} \quad \eta_{f,w} \leq \bar{v} - \tilde{\epsilon}.$$

In other words, every pair is joined by an edge independently with probability

$$p(\epsilon) = 1 - \left(1 - \Gamma^W(\bar{u} - \epsilon)\right) \cdot \left(1 - \Gamma^F(\bar{v} - \tilde{\epsilon})\right).$$

We write $\beta(n) := 2 \cdot \log(l_{\leq \bar{k}+2}(U)) / \log \frac{1}{p(\epsilon)}$, and show that

$$P\left(|\hat{B}^F_{\bar{k}}(\epsilon; U, V)| \leq \beta(n)\right) \to 1 \quad \text{as} \quad n \to \infty.$$[27]

First, observe that $\bar{W}_{\leq \bar{k}+2}(U, V)$ is the size of at least $l_{\bar{k}+2}(U)$. Among $l_{\leq \bar{k}+2}(U)$ workers in tiers up to $\bar{k} + 2$ at most $l_{\leq \bar{k}+1}(U)$ are matched with firms in tiers up to $\bar{k} + 1$. In addition, $l_{\bar{k}+2}(U) > \beta(n)$ with large n, since $\mathcal{E}_1$ holds. Therefore, with large n, we shall write

$$P\left(|\hat{B}^F_{\bar{k}}(\epsilon; U, V)| \leq \beta(n)\right) = P\left(\min\left\{|\hat{B}^F_{\bar{k}}(\epsilon; U, V)|, |\bar{W}_{\leq \bar{k}+2}(U, V)|\right\} \leq \beta(n)\right). \tag{13}$$

Let $\alpha(U, V) \times \alpha(U, V)$ be the size of a maximum balance biclique of the random graph

$$G\left(F_{\bar{k}}(U) \cup W_{\leq \bar{k}+2}(U)\ ,\ p(\epsilon)\right).$$

Since every realized $\hat{B}^F_{\bar{k}}(\epsilon; u, v) \cup \bar{W}_{\leq \bar{k}+2}(u, v)$ is a biclique, it contains a balanced biclique of the size equals to

$$\min\left\{|\hat{B}^F_{\bar{k}}(\epsilon; u, v)|\ ,\ |\bar{W}_{\leq \bar{k}+2}(u, v)|\right\}.$$

Therefore,

$$P\left(\min\left\{|\hat{B}^F_{\bar{k}}(\epsilon; U, V)|, |\bar{W}_{\leq \bar{k}+2}(U, V)|\right\} \leq \beta(n)\right) \geq P\left(\alpha(U, V) \leq \beta(n)\right). \tag{14}$$

Applying Theorem 3 to (14) and using (13),

$$P\left(|\hat{B}^F_{\bar{k}}(\epsilon; U, V)| \leq \beta(n)\right) \geq P\left(\alpha(U, V) \leq \beta(n)\right) \to 1. \tag{15}$$

Lastly, we consider random utilities U and V, in which common-values are yet realized. For every $\epsilon' > 0$,

$$\begin{aligned} P\left(\frac{|\hat{B}^F_{\bar{k}}(\epsilon; U, V)|}{n} > \epsilon'\right) &= P\left(|\hat{B}^F_{\bar{k}}(\epsilon; U, V)| > \epsilon' \cdot n\right) \\ &\leq P\left(|\hat{B}^F_{\bar{k}}(\epsilon; U, V)| > \beta(n) \mid \mathcal{E}_1, \mathcal{E}_2\right) + R_n, \quad \text{with large } n, \end{aligned}$$

[27] Note that we fixed common-values as a realization u^o and v^o such that the events $\mathcal{E}_1$ and $\mathcal{E}_2$ occur. Thus for now, the tier-structure is deterministic, and $\beta(n)$ is, in turn, a deterministic sequence.

where R_n is the probability that either $\mathcal{E}_1$ or $\mathcal{E}_2$ does not hold: i.e. $R_n \to 0$. The inequality is from the fact that $\epsilon' \cdot n > \beta(n)$ with large n. We complete the proof by applying (15). □

We also obtain the counterpart proposition of Proposition C.1 in terms of tiers defined with respect to firms' common-values.

Proposition C.1* *For each $\bar{k} = 1, 2, \ldots, K-2$, define*

$$\hat{B}^W_{\bar{k}}(\epsilon; u, v) := \left\{ w \in W_{\bar{k}}(v) | v_{\mu_F}(w) \leq v^o_{\bar{k}+2} + \bar{v} - \epsilon \right\}.$$

Then for any $\epsilon > 0$,

$$\frac{|\hat{B}^W_{\bar{k}}(\epsilon; U, V)|}{n} \xrightarrow{p} 0 \quad as \quad n \to \infty.$$

Proof. We omit the proof since it is analogous to the proof of Proposition C.1. □

For each instance $\langle F, W, u, v \rangle$ and for each $\bar{k} = 1, 2, \ldots, K$, we define

$$B^F_{\bar{k}}(\epsilon; u, v) := \{ f \in F_{\bar{k}}(u) | \Delta(f; u, v) \geq \epsilon \}.$$

Proposition C.2. *If $\bar{k} = 7, 8, \ldots, K-2$, then for any $\epsilon > 0$,*

$$\frac{|B^F_{\bar{k}}(\epsilon; U, V)|}{n} \xrightarrow{p} 0 \quad as \quad n \to \infty.$$

Proof. In Proposition C.1* with $k = 1, 2, \ldots, K-3$, we replace ϵ with

$$\epsilon_k := v^o_{k+2} - v^o_{k+3},$$

and write

$$\hat{B}^W_k(\epsilon_k; u, v) = \left\{ w \in W_k(v) | v_{\mu_F}(w) \leq v^o_{k+3} + \bar{v} \right\}.^{28}$$

Then,

$$\frac{|\hat{B}^W_k(\epsilon_k; U, V)|}{n} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty. \tag{16}$$

Note that a worker receives utility higher than $v^o_{k+3} + \bar{v}$ only by matching with a firm in (w.r.t firm) tiers up to $k + 3$.[29] Thus for $k = 5, 6, \ldots, K$,

$$\{ w \in W_{\leq k-4}(V) : \mu(w) \in F_k(V) \} \subset \bigcup_{k'=1}^{k-4} \hat{B}^W_{k'}(\epsilon_{k'}; U, V). \tag{17}$$

[28] Recall that v^o_k is a constant, defined as $v^o_k := \xi^F_{1-\frac{k}{K}}$.

[29] Recall that $f \in F_k(v)$ if and only if $v^o_k < v^o_f \leq v^o_{k-1}$. Thus, if $f \in F_{>k+3}(v)$ then $v^o_f \leq v^o_{k+3}$.

If event $\mathcal{E}_3$ holds, we can translate (17) into an expression with tiers w.r.t workers' common-values. That is, for $k = 7, 8, \dots, K$,

$$\begin{aligned}\{w \in W_{\leq k-6}(U) : \mu_F(w) \in F_k(U)\} &\subset \bigcup_{k'=k-1}^{k+1} \{w \in W_{\leq k-6}(U) : \mu_F(w) \in F_{k'}(V)\} \\ &\subset \bigcup_{k'=k-1}^{k+1} \{w \in W_{\leq k-5}(V) : \mu_F(w) \in F_{k'}(V)\} \\ &\subset \bigcup_{k'=k-1}^{k+1} \{w \in W_{\leq k'-4}(V) : \mu_F(w) \in F_{k'}(V)\}\end{aligned}$$

where the first and second inequalities are from $\mathcal{E}_3$.

By applying (17), we obtain

$$\{w \in W_{\leq k-6}(U) : \mu_F(w) \in F_k(U)\} \subset \bigcup_{k'=1}^{k-3} \hat{B}^W_{k'}(\epsilon_{k'}; U, V).$$

It follows that

$$\frac{|\{f \in F_k(U) : \mu_F(f) \in W_{\leq k-6}(U)\}|}{n} \xrightarrow{p} 0, \tag{18}$$

because for every $\epsilon > 0$,

$$P\left(\frac{|\{f \in F_k(U) : \mu_F(f) \in W_{\leq k-6}(U)\}|}{n} > \epsilon\right) \leq P\left(\sum_{k'=1}^{k-3} \frac{|\hat{B}^W_{k'}(\epsilon_{k'}; U, V)|}{n} > \epsilon\right) + R_n,$$

where R_n is the probability that $\mathcal{E}_3$ does not hold: i.e. $R_n \to 0$. The right hand side converges to 0 by (16).

We complete the proof of Proposition C.2 by proving the following claim. Proposition C.1 and (18) show that the normalized sizes of two sets on the right hand side of (19) converge to 0 in probability.

Claim C.1. *For $\bar{k} = 7, 8, \dots, K-2$ and each instance $\langle F, W, u, v\rangle$,*

$$B^F_{\bar{k}}(\epsilon; u, v) \subset \hat{B}^F_{\bar{k}}(\epsilon/9; u, v) \cup \left\{f \in F_{\bar{k}}(u) | \mu_F(f) \in W_{\leq \bar{k}-6}(u)\right\}. \tag{19}$$

Proof of Claim C.1. If a firm $f \in F_{\bar{k}}(u)$ is *not* in $\hat{B}^F_{\bar{k}}(\epsilon/9; u, v)$, then

$$u_{\mu_W}(f) > u^o_{\bar{k}+2} + \bar{u} - \epsilon/9,$$

and if the firm f is *not* in $\left\{f \in F_{\bar{k}}(u) | \mu_F(f) \in W_{\leq \bar{k}-6}(u)\right\}$, then

$$u_{\mu_F}(f) \leq u^o_{\bar{k}-6} + \bar{u}.$$

Therefore, using (10) we obtain

$$u_{\mu_F}(f) - u_{\mu_W}(f) \leq u^o_{k-6} - u^o_{k+2} + \epsilon/9 < \epsilon,$$

and thus f is *not* in $B^F_k(\epsilon; u, v)$.

$\square$

$\square$

Lastly, we complete the proof of Theorem 1 by the following inequalities.

$$\begin{aligned}
P\left(\frac{|B^F(\epsilon; U, V)|}{n} > \frac{9}{K}\right) &= P\left(\sum_{1\leq k\leq K} \frac{|B^F_k(\epsilon; U, V)|}{n} > \frac{9}{K}\right) \\
&< P\left(\sum_{7\leq k\leq K-2} \frac{|B^F_k(\epsilon; U, V)|}{n} + \sum_{k=1,\ldots,6,K-1,K} \frac{l_k(U)}{n} > \frac{9}{K}\right).
\end{aligned}$$

The last probability converges to 0. For each $k = 7, \ldots, K-2$, the proportion $\frac{|B^F_k(\epsilon;U,V)|}{n}$ converges to 0 in probability (Proposition C.2). For each $k = 1, \ldots, 6, K-1, K$, the proportion $\frac{l_k(U)}{n}$ converges to $\frac{1}{K}$ in probability by the (weak) law of large numbers.

D Proof of Theorem 4

For each $\epsilon > 0$, we first define

$$B^F_E(\epsilon; u, v) := F \backslash A^F_E(\epsilon; u, v) = \{f \in F \mid \Delta_E(f; u, v) \geq \epsilon\},$$

and show that

$$E\left[\frac{|B^F_E(\epsilon; U, V)|}{n}\right] \to 0 \quad \text{as} \quad n \to \infty.$$

For each n, let $f_n \in F_n$ and consider the resulting sequence $\langle f_n \rangle_{n=1}^{\infty}$. For any $\epsilon > 0$,

$$\begin{aligned}
E\left[\frac{|B^F_E(\epsilon; U, V)|}{n}\right] &= E\left[\mathbf{1}\{f_n \in B^F_E(\epsilon; U, V)\}\right] \\
&= P\left(\Delta_E(f_n; U, V) \geq \epsilon\right).
\end{aligned}$$

Thus if $\Delta_E(f_n; U, V) \xrightarrow{p} 0$, then for every ϵ, $\frac{|B^F_E(\epsilon,U,V)|}{n}$ converges to zero in mean, thereby completing the proof.

Claim D.1.

$$\Delta_E(f_n; U, V) \xrightarrow{p} 0, \quad \textit{as} \quad n \to \infty.$$

Proof. For every $\epsilon > 0$,

$$P\left(\Delta(f_n; U, V) \geq \epsilon\right) = E\left[\mathbf{1}\{\Delta(f_n; U, V) \geq \epsilon\}\right] \\ = E\left[\frac{|F \backslash A^F(\epsilon; U, V)|}{n}\right].$$

The last term converges to 0 by Theorem 1, and thus $\Delta(f_n; U, V) \xrightarrow{p} 0$.

Let $\bar{u}^o$ and $\bar{u}$ be upper bounds of common-value distribution and private-value distribution of workers, respectively. Then, $\Delta(f_n; U, V)$ is bounded above by $\lambda\, \bar{u}^o + (1 - \lambda)\bar{u}$ with probability 1. We obtain by Theorem A.2 that

$$\lim_{n\to\infty} E[\Delta_E(f_n; U, V)] = \lim_{n\to\infty} E\left[E\left[\Delta(f_n; U, V)|\Pi_{f_n}\right]\right] = \lim_{n\to\infty} E\left[\Delta(f_n; U, V)\right] = 0.$$

The Claim D.1 follows by Theorem A.1. □

E Additional Simulations on the Proportion of Unmatched Agents

The simulation results in Section 1.4 show that the short preference condition assumed in Roth and Peranson (1999), Immorlica and Mahdian (2005), and Kojima and Pathak (2009) may leave most agents in a large market unmatched in stable matchings. It is worth noting that random preferences in the previous simulations were generated by the setup of our model, rather than the previous studies' model. That is, the previous simulations do not directly represent features of previous models. In this section, we show the increasing proportions of unmatched agents with simulations based on the previous studies' model.

Let L be the maximum number of firms that each worker considers acceptable. We generate random preferences following the previous model, in particular Immorlica and Mahdian (2005). Immorlica and Mahdian studied one-to-one matching markets with generally distributed random preferences. For each market size n, a market is given two underlying distributions, one for firms and the other for workers, called *popularity distributions*.[30] A worker's preference list is constructed by sequentially sampling L firms from the popularity distribution without replacement. The firm chosen first is the most preferred, and the next chosen firm becomes the second most preferred. We similarly construct firms' preferences, except that firms' preferences are of length n: i.e., all workers are acceptable.

We use two classes of popularity distributions.

1. Normalized geometric distribution

[30] Immorlica and Mahdian (2005) construct random preferences only for workers: firms' preferences are arbitrarily given. In our simulation, we also generate firms' preferences randomly, rather than assuming particular preferences. Accordingly, we can measure the general likely proportions of unmatched agents.

For each market size n, we define the normalized geometric distribution as:

$$\text{PDF}: p_k = \frac{(1-q)^k}{\sum_{k'=1}^{n}(1-q)^{k'}}, \quad (0 \le q < 1,\ k = 1, 2, \ldots, n).$$

Consider a pair of firms, f_{k_1} and f_{k_2} ($k_1 < k_2 \le n$). For each worker, the probability of choosing f_{k_1} before f_{k_2}, conditioned on at least one of the firms chosen, equals to

$$\frac{(1-q)^{k_1}}{(1-q)^{k_1} + (1-q)^{k_2}} = \frac{1}{1 + (1-q)^{k_2 - k_1}}.$$

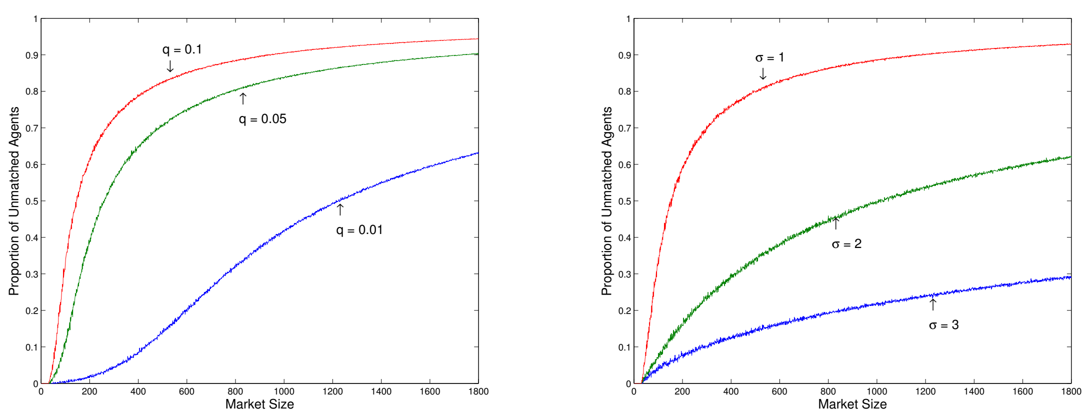
which is independent of the market size n. If $q = 0$, we have the uniform popularity distribution over firms, so all firms have an equal chance of being chosen before another. As q becomes close to 1, more popular firms have higher chances of being chosen before other firms, which generates a commonality of preferences among workers.

2. Normalized log-normal distribution

   Let $F(\,\cdot\,; \mu, \sigma)$ be the cumulative distribution function of a log-normal distribution. For each market size n, we define the normalized log-normal distribution as:

$$\text{PDF}: p_k = \frac{F(\,k\,; \mu, \sigma) - F(\,k-1\,; \mu, \sigma)}{F(\,n\,; \mu, \sigma)}, \quad (\mu, \sigma \in \mathbb{R},\ k = 1, 2, \ldots, n).$$

   For each μ, as σ increases, firms have similar probabilities to be chosen. This generates a weaker commonality of preferences among workers.



(a) Normalized geometric distribution (b) Normalized log-normal distribution

Figure 3: Proportions of unmatched agents in stable matchings.

Figure 3 shows that the proportion of unmatched agents in stable matchings increases as a market becomes large. Each graph represents the proportion of unmatched agents, when workers consider 30 most preferred firms acceptable. The proportions are averaged over 10 repetitions.

References

ABDULKADIROGLU, A., P. PATHAK, AND A. ROTH (2009): "Strategy-proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match," *The American Economic Review*, 99(5), 1954–1978.

ABDULKADIROGLU, A., P. PATHAK, A. ROTH, AND T. SÖNMEZ (2006): "Changing the Boston school choice mechanism," .

ABDULKADIROGLU, A., AND T. SÖNMEZ (2003): "School choice: A mechanism design approach," *The American Economic Review*, 93(3), 729–747.

ALCALDE, J., AND S. BARBERÀ (1994): "Top dominance and the possibility of strategy-proof stable solutions to matching problems," *Economic theory*, 4(3), 417–435.

ASHLAGI, I., M. BRAVERMAN, AND A. HASSIDIM (2011): "Stability in Large Matching Markets with Complementarities," Discussion paper, Mimeo.

AZEVEDO, E. (2010): "Imperfect Competition in Two-Sided Matching Markets," Discussion paper, working paper, Harvard University.

AZEVEDO, E., AND J. LESHNO (2011): "The College Admissions Problem With a Continuum of Students," Discussion paper, Mimeo.

BACCARA, M., A. COLLARD-WEXLER, L. FELLI, AND L. YARIV (2010): "Child adoption matching: preferences for gender and race," .

BERNAL, R., L. HU, C. MORIGUCHI, AND E. NAGYPAL (2007): "Child Adoption in the United States: Historical Trends and the Determinants of Adoption Demand and Supply, 1951-2002," .

BOGOMOLNAIA, A., AND H. MOULIN (2001): "A new solution to the random assignment problem," *Journal of Economic Theory*, 100(2), 295–328.

BULOW, J., AND J. LEVIN (2006): "Matching and Price Competition," *American Economic Review*, 96(3), 652–668.

CHE, Y., AND F. KOJIMA (2010): "Asymptotic equivalence of probabilistic serial and random priority mechanisms," *Econometrica*, 78(5), 1625–1672.

CHOO, E., AND A. SIOW (2006): "Who marries whom and why," *Journal of Political Economy*, 114(1), 175.

CLARK, S. (2006): "The uniqueness of stable matchings," *The BE Journal of Theoretical Economics*, 6(1), 8.

Coles, P., J. Cawley, P. Levine, M. Niederle, A. Roth, and J. Siegfried (2010): "The Job Market for New Economists: A Market Design Perspective," *Journal of Economic Perspectives*, 24(4), 187–206.

Crawford, V. (2008): "The flexible-salary match: a proposal to increase the salary flexibility of the national resident matching program," *Journal of Economic Behavior & Organization*, 66(2), 149–160.

Dawande, M., P. Keskinocak, J. Swaminathan, and S. Tayur (2001): "On bipartite and multipartite clique problems," *Journal of Algorithms*, 41(2), 388–403.

Demange, G., D. Gale, and M. Sotomayor (1987): "A further note on the stable matching problem," *Discrete Applied Mathematics*, 16, 217–222.

Dubins, L., and D. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *The American Mathematical Monthly*, 88(7), 485–494.

Eeckhout, J. (2000): "On the uniqueness of stable marriage matchings," *Economics Letters*, 69(1), 1–8.

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

Halaburda, H. (2010): "Unravelling in two-sided matching markets and similarity of preferences," *Games and Economic Behavior*, 69(2), 365–393.

Hitsch, G., A. Hortaçsu, and D. Ariely (2010): "What makes you click?Mate preferences in online dating," *Quantitative Marketing and Economics*, 8(4), 393–427.

Immorlica, N., and M. Mahdian (2005): "Marriage, honesty, and stability," in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pp. 53–62. Society for Industrial and Applied Mathematics.

Knuth, D. (1976): *Mariages stables*. Les Presse De L'Universite De Montreal.

Kojima, F., and M. Manea (2010): "Incentives in the probabilistic serial mechanism," *Journal of Economic Theory*, 145(1), 106–123.

Kojima, F., and P. Pathak (2009): "Incentives and stability in large two-sided matching markets," *The American Economic Review*, 99(3), 608–627.

Kojima, F., P. Pathak, and A. Roth (2010): "Matching with Couples: Stability and Incentives in Large Markets," *NBER Working Papers*.

Liu, Q., and M. Pycia (2011): "Ordinal Efficiency, Fairness, and Incentives in Large Markets," .

McKinney, C., M. Niederle, and A. Roth (2003): "The collapse of a medical clearinghouse (and why such failures are rare)," .

McVitie, D., and L. Wilson (1970): "Stable marriage assignment for unequal sets," *BIT Numerical Mathematics*, 10(3), 295–309.

Niederle, M., and A. Roth (2005): "The gastroenterology fellowship market: Should there be a match?," *The American economic review*, 95(2), 372–375.

Pathak, P., and T. Sonmez (2008): "Leveling the playing field: Sincere and sophisticated players in the Boston mechanism," *The American Economic Review*, 98(4), 1636–1652.

Pittel, B. (1989): "The Average Number of Stable Matchings," *SIAM Journal on Discrete Mathematics*, 2(4), 530–549.

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

——— (1984): "The evolution of the labor market for medical interns and residents: a case study in game theory," *The Journal of Political Economy*, 92(6), 991–1016.

——— (2002): "The economist as engineer: Game theory, experimentation, and computation as tools for design economics," *Econometrica*, 70(4), 1341–1378.

——— (2008): "Deferred acceptance algorithms: History, theory, practice, and open questions," *International Journal of Game Theory*, 36(3), 537–569.

Roth, A., and E. Peranson (1999): "VThe Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *V American Economic Review*, 89(748), 80.

Roth, A., and U. Rothblum (1999): "Truncation strategies in matching marketsin search of advice for participants," *Econometrica*, 67(1), 21–43.

Roth, A., and M. Sotomayor (1990): *Two-sided matching*. Cambridge Univ. Pr.

Roth, A., and J. Vande Vate (1991): "Incentives in two-sided matching with random stable mechanisms," *Economic theory*, 1(1), 31–44.

Roth, A., and X. Xing (1994): "Jumping the gun: imperfections and institutions related to the timing of market transactions," *The American Economic Review*, 84(4), 992–1044.

Samet, D. (2011): "Matching of like rank and the size of the core in the marriage problem," Discussion paper, Mimeo.

Serfling, R. (1980): *Approximation theorems of mathematical statistics*, vol. 371. Wiley New York.

Sönmez, T. (1999): "Strategy-proofness and Essentially Single-valued Cores," *Econometrica*, 67(3), 677–689.

Sönmez, T., and T. Switzer (2011): "Matching with (Branch-of-Choice) Contracts at United States Military Academy," *Boston College Working Papers in Economics*.