BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2017-4

Stability and Efficiency in Decentralized Two-Sided Markets with Weak Preferences

by Radoslav S. Raykov

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2017-4

February 2017

Stability and Efficiency in Decentralized Two-Sided Markets with Weak Preferences

by

Radoslav S. Raykov

Financial Stability Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
rraykov@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

Utku Ünver and Hideo Konishi kindly provided advice for an early version of this paper. The author wishes to thank the following individuals: Yeon-Koo Che, Aytek Erdil, Fuhito Kojima, Hanna Halaburda, Sonya Kiselgof, Sofia Priazhkina, Marilda Sotomayor, participants in the BWGT and SED conferences, and colleagues at the Bank of Canada for feedback on previous versions of this work. The views and opinions expressed in this article are the author's and do not necessarily represent the views of the Bank of Canada.

Abstract

Many decentralized markets are able to attain a stable outcome despite the absence of a central authority (Roth and Vande Vate, 1990). A stable matching, however, need not be efficient if preferences are weak. This raises the question whether a decentralized market with weak preferences can attain Pareto efficiency in the absence of a central matchmaker. I show that when agent tastes are independent, the random stable match in a large-enough market is asymptotically Pareto efficient even with weak preferences. In fact, even moderate-sized markets can attain good efficiency levels. The average fraction of agents who can Pareto improve is below 10% in a market of size n = 79 when one side of the market has weak preferences; when both sides have weak preferences, the inefficiency falls below 10% for n > 158. This implies that approximate Pareto efficiency is attainable in a decentralized market even in the absence of a central matchmaker.

Bank topics: Economic models
JEL codes: C78; D61

Résumé

Beaucoup de marchés décentralisés peuvent atteindre un état stable malgré l'absence d'autorité centrale (Roth et Vande Vate, 1990). Un jumelage stable, toutefois, n'a pas besoin d'être efficient si les préférences sont peu marquées. On peut dès lors se demander si un marché décentralisé caractérisé par des préférences peu marquées peut atteindre une efficience parétienne en l'absence d'un planificateur. L'auteur montre que, lorsque les goûts des agents économiques sont indépendants, le jumelage aléatoire stable dans un marché suffisamment important est efficient asymptotiquement au sens de Pareto même si les préférences sont peu marquées. En fait, les marchés de taille modeste peuvent eux aussi atteindre un degré d'efficience raisonnable. La proportion d'agents dont la situation est susceptible de s'améliorer au sens de Pareto est, en moyenne, inférieure à 10 % dans un marché de taille n = 79 lorsque les offreurs ou les demandeurs du marché ont des préférences peu marquées; si les deux groupes ont des préférences peu marquées, l'inefficience tombe au-dessous de 10 % pour n > 158. Ces résultats portent à croire qu'un marché décentralisé peut se rapprocher d'une efficience parétienne même en l'absence de planificateur.

Sujet : Modèles économiques
Codes JEL : C78, D61

Non-Technical Summary

To operate properly, many markets rely on matching economic agents such as buyers with sellers, resident physicians with hospitals, and kidney donors with kidney patients, to name a few. The notion of stability (the tendency for a matching not to dissolve) plays a central role in such markets. Stability has been studied extensively in the literature, and has been shown to also imply Pareto efficiency when agent preferences are strict.

However, when preferences admit indifferences, not every stable matching need be Pareto efficient. Instead, some agents may benefit from a reassignment without making anyone else worse off. In stable but inefficient matchings, stability creates perverse incentives for inefficiently matched agents to stay together, so that the match has no natural tendency to dissolve or improve its efficiency without a central matchmaker. Since decentralized markets do not have a central coordinating authority, this raises the question of how a decentralized market with weak preferences can attain Pareto efficiency.

I approach this question in the spirit of several newer papers showing that certain shortcomings of stable two-sided matchings tend to improve in large markets (Kojima and Pathak, 2009; Kojima and Manea, 2010). In this paper, I demonstrate that with weak preferences, a random stable matching will be approximately Pareto efficient in a large market. Specifically, I show that the random stable mechanism, which selects a matching at random from the uniform distribution over all stable matchings, attains approximate Pareto efficiency in large decentralized markets.

Since the result is asymptotic, it is natural to ask how large a market is large enough to result in acceptable efficiency levels. In practice, even moderate-sized markets attain good efficiency using this metric. The market size necessary to reduce the proportion of agents admitting Pareto-improvement cycles to less than 10% is only n = 79 when one side of the market admits weak preferences, and n = 158 when both sides admit weak preferences. The implied efficiency level above 90% is encouraging for one-to-one matchings. A caveat is that these results do not generalize to many-to-one matching. Rather than offering a “cure-all solution,” this paper provides an asymptotic upper bound for the inefficiency and shows that even in relatively small markets, this inefficiency is small.

1 Introduction

Many two-sided matching markets operate in decentralized settings where agents meet randomly and match in the absence of a central authority. Roth and Vande Vate (1990) first showed that despite their decentralized nature, such markets can attain stable outcomes by following random processes such as *random paths to stability*, where random blocking pairs are sequentially resolved over time. The salient characteristic of such markets is that despite their decentralized nature, a stable matching can nonetheless be attained.

However, when preferences admit indifferences, a stable matching need not be Pareto efficient (Roth and Sotomayor, 1990). In stable but inefficient matchings, stability creates perverse incentives for inefficiently matched agents to stay together, so that the match has no tendency to dissolve or improve its efficiency over time. Unlike random paths to stability, there is no natural process by which an inefficient stable matching could evolve to a more Pareto efficient one without a central matchmaker. This raises the question of whether a decentralized market with weak preferences can attain Pareto efficiency in the absence of a central authority.

I approach this question in the spirit of several newer papers showing that certain shortcomings of stable matchings can improve in large markets (Kojima and Pathak, 2009; Kojima and Manea, 2010). I demonstrate that with weak preferences, a random stable matching will be approximately Pareto efficient in a large market with non-transferable utility. Specifically, I show that the *random stable mechanism*, which selects a matching at random from the uniform distribution over all stable matchings, attains approximate Pareto efficiency in large decentralized markets. While random paths may be closer to the actual matching process, the random stable mechanism is a clean benchmark that provides better conceptual insight about how heterogeneity helps correct Pareto inefficiency in large markets.

The metric I use to gauge Pareto inefficiency is the proportion of agents who benefit from Pareto-improving exchanges known as Pareto-improvement cycles and chains (Erdil and Ergin, 2015). Since the remaining agents cannot improve from a reassignment, this metric is suitably conservative. I show that the proportion of agents in a random stable match who can Pareto improve (and therefore admit Pareto-improvement cycles or chains) is vanishing in a large market.

The first necessary condition for this result is that preferences must be independent, as in a decentralized marketplace where agents do not systematically influence each other's tastes;

and secondly, the market must be sufficiently large, as measured by the number of agents on each side of the market. To reflect this setting, similar to Kojima and Pathak (2009), I consider a sequence of *random markets* — instances of the marriage problem of increasing size, where agent preferences are drawn from a probability distribution.

Since the result is asymptotic, it is natural to ask how large a market is large enough to result in acceptable efficiency levels. In practice, even modest-sized markets attain good efficiency using this metric. The market size necessary to reduce the proportion of agents admitting Pareto-improvement cycles to less than 10% is only $n = 79$ when one side of the market admits weak preferences, and $n = 158$ when both sides admit weak preferences. The implied efficiency level above 90% is encouraging for one-to-one matchings, although the results do not generalize to many-to-one matching. Rather than offering a "cure-all," this paper provides an asymptotic upper bound and shows that even in relatively small markets, the inefficiency is small.

My strategy is to first identify inefficiently matched agents using the concept of Pareto-improvement cycles and chains (Erdil and Ergin, 2015). A Pareto-improvement cycle (or chain, when unmatched agents exist) is a set of agents who can switch partners amongst themselves so that at least one agent in the set is strictly better off, while nobody else is worse off; it can be shown that a matching is Pareto efficient if and only if it admits neither Pareto-improvement cycles nor Pareto-improvement chains. Such cycles and chains, therefore, are natural identifiers for potentially inefficient matchings. My strategy is to develop an asymptotic approximation for the probability of Pareto-improvement cycles (chains) and show that this probability converges to zero in a large market. Further, I show that the same is true of the expected proportion of agents who can Pareto improve.

The smallest component of a cycle consists of a 2-tuple of agents from the same side of the market (two men or two women). The statement "a particular man admits a cycle" is meaningless unless it is also specified with what other man, since the latter man's partner must also be willing to switch. For this reason, I focus on the asymptotic probability that an arbitrary 2-tuple of men (women) admits a cycle. As an intermediate result, I show that a cycle cannot arise between two men in a stable matching unless one of their mates is indifferent between them, and then I seek to approximate this probability.

To gauge the probability that a given 2-tuple admits a cycle, one must know both how frequently a particular man is matched to a particular woman, and how frequently the latter woman is indifferent between the two men in the 2-tuple. Generally, these two events are statistically dependent, because the matching produced by the mechanism depends on the preferences that have been fed to it. However, I show that in a large market, the allocation obtained by the random stable mechanism becomes independent of indifferences, permitting one to disentangle the influence of preferences on the resulting matching from that of the mechanism. This allows one to translate the maximum probability of a cycle in terms of the probability of a pairwise indifference in a random weak preference, which I quantify for large n using combinatorial methods.

This asymptotic decoupling involves a few technical steps that are of interest on their own. First, I show that the random stable mechanism is *symmetric* in matching each man to each woman with equal likelihood when agent preferences are uniform, independent draws. However, conditioning on a pairwise indifference (the necessary condition for cycles) violates symmetry, because the mechanism's output links back to the asymmetric preferences that have been fed to it. Thus the allocation depends statistically on who can form cycles with whom; the two events are not independent. One technical contribution of the paper is to show that this statistical dependence vanishes and symmetry is restored as $n \to \infty$. This allows me to approximate the probability of a 2-tuple admitting a cycle simply with the probability of the corresponding pairwise indifference.

Mechanisms other than the random stable mechanism generally do not share this independence property, because the matching correlates with the agent preferences used as an input. Thus, anything that upsets the symmetry of the mechanism, or the symmetry, uniformity and mutual independence of agent preferences, might also upset the asymptotic independence property that allows one to make inferences about the behavior of cycles. Nonetheless, I consider the random stable mechanism with a very broad class of non-uniform preferences and show that, as long they remain symmetric and independent, the same or stronger results hold. The reason for that is that the paper's main insight — that large-market heterogeneity makes it harder for agent preferences to line up in the concrete way required to form a cycle — is conceptual rather than technical.

The rest of the paper is organized as follows. Section 2 provides a brief literature review, and section 3 the model and main results. Section 4 discusses the results, and section 5 concludes. Proofs are relegated to the Appendix.

2 Related Literature

Roth and Vande Vate (1990) first showed that an arbitrary initial matching can evolve into a stable matching with probability 1 if randomly chosen blocking pairs are resolved over time. This result provides a strong justification why some decentralized markets are able to function smoothly in the absence of a central coordinating body.[1]

The random stable outcome can be path-dependent (depending on which particular random sequence of blocking pairs is resolved), or path-independent, such as in a random selection among stable matchings. In this paper, I focus on the path-independent approach as implemented by the random stable mechanism (RSM) — the mechanism that selects a stable matching at random from the uniform distribution over all stable matchings. While the path-dependent approach may be closer to how markets work, as a clean benchmark case, the RSM provides better conceptual insight about how Pareto inefficiency corrects itself in large markets.

When agents have strict preferences, stability already implies efficiency, so a stable match is also Pareto efficient. With weak preferences, however, stability no longer guarantees Pareto efficiency. To see this, consider a simple marriage problem with three men and three women, where men have the strict preferences P_i and women have the weak preferences R_i given below,

$$P_{m_1} : w_1,\ w_2,\ w_3,\ m_1; \quad R_{w_1} : \{m_1,\ m_2,\ m_3\},\ w_1$$
$$P_{m_2} : w_1,\ w_3,\ w_2,\ m_2; \quad R_{w_2} : \{m_1,\ m_2,\ m_3\},\ w_2$$
$$P_{m_3} : w_2,\ w_1,\ w_3,\ m_3; \quad R_{w_3} : \{m_1,\ m_2,\ m_3\},\ w_3$$

where potential mates are listed in a descending order of preference and braces indicate mates that are equivalent in terms of utility. Now consider the following two stable matchings, μ_1 and μ_2:

$$\mu_1 = \begin{pmatrix} m_1 & m_2 & m_3 \\ w_2 & w_3 & w_1 \end{pmatrix} \qquad \mu_2 = \begin{pmatrix} m_1 & m_2 & m_3 \\ w_1 & w_3 & w_2 \end{pmatrix}.$$

The allocation μ_1 is stable, but inefficient, because men m_1 and m_3 can switch partners and obtain their respective first choices without anyone being worse off and without upsetting stability, thus arriving at the stable matching μ_2. Hence, stability not only does not imply Pareto efficiency, but in addition creates bad incentives for inefficiently matched agents to

[1] Roth and Vande Vate (1990) use entry-level labor markets with non-transferable utility as an example.

stay together, since inefficient stable matchings have no natural tendency to dissolve or Pareto improve over time.

Such matchings therefore pose not only theoretical but also practical concerns, such as in high-school admissions. For example, Abdulkadiroğlu, Pathak and Roth (2009) show that inefficient stable matchings in New York City high-school admission result in more than 6,800 students per year receiving a Pareto inefficient placement. For this reason, improving the efficiency of stable matchings in markets with weak preferences remains in the focus of recent work (Erdil and Ergin, 2008).

However, several new studies have shown that certain finite-market shortcomings of a mechanism can vanish when the market is large. For example, Kojima and Pathak (2009) find that incentives to manipulate the student-optimal mechanism greatly diminish in large markets, so that truth-telling becomes an ε-Nash equilibrium. Kojima and Manea (2010) obtain a similar result for the probabilistic serial mechanism, and Che and Kojima (2010) demonstrate that even seemingly unrelated mechanisms can converge to the same outcome in a large market. These results show that even well-studied mechanisms can behave surprisingly in large allocation problems, allowing the mechanism designer to sometimes improve on existing shortcomings.

Roth and Peranson (1999) and Immorlica and Mahdian (2005) first introduced two-sided matching in a random environment where agent preferences are drawn from a probability distribution. This technique allows a parsimonious representation of decentralized markets whose heterogeneity increases with market size, and was extended to many-to-one matching by Kojima and Pathak (2009), who reintroduced it under the name *random market*. It conveniently allows the mechanism designer to predict the average market outcome with increasing accuracy as the number of agents grows. For instance, Ashlagi et al. (2015) use the random market device to show that in an unbalanced marriage market, the core collapses to a single stable matching with a probability close to 1. In contrast to Ashlagi et al. (2015), the marriage problem considered here generally has a much larger, non-singleton core — first, because the market is balanced, and second, because weak preferences increase the number of stable matchings. In my paper, random preferences are important insofar as they help generate heterogeneity, which increases with market size and makes it harder for preferences to line up in the way required to form a cycle.

3 Model and Results

A *marriage problem*[2] is a triple $(M, W, R)_n$ consisting of two sets of agents, M and W, each of cardinality n, whose members must form pairs containing one agent of each set according to their collection of preferences $R = \{R_m \cup R_w,\ \ m \in M,\ w \in W\}$. For concreteness, the members of M are typically dubbed "men" and those of W, "women." Utility is ordinal and non-transferable, so preferences can be thought of as rank lists in which each agent lists all possible mates in order of preference. It is standard to denote the weak preference relation "at least as good as" with the letter R indexed by the identity of the agent: for example, aR_ib denotes "agent i weakly prefers a to b". The corresponding strict preference relation is denoted with P_i, and the indifference relation with $\sim_i$. I will often refer to the number of agents n on each side of the market as simply *the market size*.

Given agent preferences, one can use a deterministic or random procedure (mechanism) to allocate partners. An allocation specifying exactly one partner for every agent is called a *matching*, denoted with μ or μ_n (when desiring to explicitly reference the size of the market).[3] Naturally, different preferences can result in different matchings. The notation $\mu(i)$ denotes the partner assigned to agent i by the matching μ. A one-to-one correspondence between k men and k women (for $k < n$) is called a *k-assignment*, so as to distinguish it from a matching (since the same assignment can occur in different matchings assigning the remaining partners differently).

A matching μ is *Pareto efficient* when there exists no other matching ν under which everybody is at least weakly better off and at least one agent is strictly better off in comparison to μ. A matching μ is *individually rational* if every agent weakly prefers his or her match under μ compared with staying single. Individual rationality is closely linked to the concept of a blocking individual. The matching μ is *blocked by an individual* i if i strictly prefers remaining single to the partner assigned by μ. Thus, a matching is individually rational if it is not blocked by any individual, thereby guaranteeing that the allocation is not, in some sense, undesirable. A notion stronger than individual rationality is that of stability, which involves a more extensive concept of blocking. A matching μ is *blocked by a pair* (m, w) if m and w are not assigned to each other, yet both would strictly prefer to be together compared with the partners assigned to them by μ. When there are no blocking pairs or blocking

[2]The italicized terms in this section are defined more rigorously in Roth and Sotomayor (1990).
[3]If an agent remains single, by convention, the agent is matched to himself (herself).

individuals, we say that the matching is *stable*.[4] The stability of the match — or the lack thereof — has been successfully linked to the performance of many two-sided markets, such as the market for resident physicians (Roth, 1984; Roth, 1991). Roth and Vande Vate (1990) first formally demonstrated that permitting random blocking pairs to match over time can transform an arbitrary matching into a stable one with probability 1.

My goal is to arrive at a parsimonious representation of a two-sided matching market with decentralized decision-making, where a large number of heterogeneous agents meet and match randomly, without exerting a systematic influence on each other's tastes.[5] A convenient way to represent this heterogeneity is to treat individual preferences as independent random draws from the preference distribution.

I consider several classes of preference-generating processes attaining this goal. *For simplicity, I assume that only women's random preferences are weak*; the case where both sides of the market admit weak preferences is addressed in the Appendix.

Definition 1. A *preference-generating process* (PGP) $\mathscr{P}^i$ for an agent $i \in M \cup W$ consists of a set of preferences from which i is allowed to choose, together with a distribution over allowable preferences. A preference-generating process for all agents $\mathscr{P}$ is a collection of processes $\mathscr{P}^i$ specifying the PGP for each agent $i \in M \cup W$.

The *default preference-generating process*, $\mathscr{P}_0$, assumes that agent preferences are drawn randomly and mutually independently from the uniform distribution over all possible rankings of agents of the opposite sex, so it is used as a benchmark.

Definition 2. The *default preference-generating process* $\mathscr{P}_0$ is defined as follows:

a) $\mathscr{P}_0^m$: If m is a man, m's realized preference P_m is drawn from the uniform probability distribution over all possible strict preferences.

b) $\mathscr{P}_0^w$: If w is a woman, w's realized preference R_w is drawn from the uniform probability distribution over all possible weak preferences.

[4] One can also define a stricter notion of stability. A matching is *strictly stable* if it is individually rational and there is no pair (m, w) such that either $wR_m\mu(m)$ and $mP_w\mu(w)$ hold together, or $wP_m\mu(m)$ and $mR_w\mu(w)$ hold together. Throughout the paper, I use the regular notion of stability, sometimes also called "weak stability" (as in Gusfield and Irving, 1989); it is trivial to see that all strictly stable matchings are Pareto efficient.

[5] Preferences are formed before the matching process begins and are not state-dependent.

c) Individual agents' draws are mutually independent.

For comparison, section 4 also considers a broad class of symmetric, non-uniform PGPs $\mathscr{P}^*$, where pairwise indifferences vanish at faster rates. For benchmarking, I also consider a restricted-preference process $\mathscr{P}_1$, defined in the Appendix, where the uniform choice is restricted only to preferences where two fixed opposite-sex agents are equivalent (the necessary condition for a cycle). This benchmarking yields valuable insight about the symmetry of the outcome.

When preferences permit indifferences, the number of possible weak preferences over the opposite sex is substantially larger than with strict preferences. For example, there are only 6 ways to rank strictly three men m_1, m_2 and m_3, but there are 13 ways to order them weakly, because the order and position of indifference classes matter. In addition to the 6 strict rankings, one can also create the following 7 rankings containing at least one indifference class by grouping agents into indifference classes and varying their order:

$$m_1, \{m_2, m_3\}; \quad \{m_2, m_3\}, m_1; \qquad m_2, \{m_1, m_3\}; \quad \{m_1, m_3\}, m_2;$$
$$m_3, \{m_1, m_2\}; \quad \{m_1, m_2\}, m_3; \qquad \{m_1, m_2, m_3\}.$$

These additional seven rankings bring the total number of weak preferences over 3 men to a total of 13. As the number of agents n grows, it can be shown that the number of weak rankings grows substantially faster than $n!$. Thus, simply keeping track of the number of possible weak preferences becomes a challenge even for market sizes as low as 10 or 15. For example, the possible number of weak preferences over 10 agents is 102,247,563 and, over 15 agents, equals 230,283,190,977,853.

To facilitate the orderly bookkeeping of weak preferences, it is convenient to think of them as *ordered partitions* of the set of agents from the opposite side of the market. Under this scheme, generating a weak preference involves grouping opposite-side agents into blocks of various sizes corresponding to indifference classes, and arranging those blocks in order of preference. For example, the following ordered partition of the set of eight men $\{m_1, ..., m_8\}$

$$m_7 \mid m_2\ m_3\ m_5 \mid m_4 \mid m_6\ m_1 \mid m_8$$

corresponds to the weak preference

$$m_7, \{m_2\ m_3\ m_5\}, m_4, \{m_6\ m_1\}, m_8.$$

Thus, drawing a random preference from the weak-preference distribution is equivalent to building a random ordered partition of the set of agents from the opposite market side. I use the equivalence between weak preferences and ordered partitions in order to derive an expression for the total number of weak preferences and the relative frequency of pairwise indifferences. As explained in the introduction, the ability to link the relative frequency of Pareto-improvement cycles and chains to that of pairwise indifferences in the underlying preference profile is essential for making the problem tractable.

Even with this shortcut, however, it turns out that there is no closed-form expression for the total number of weak preferences (ordered partitions) T_n of a set with cardinality n. Instead, using combinatorial methods, I derive a recurrence relation for the total number of weak preferences. A closely related number of interest is the total number $\tilde{T}_n$ of cases where two fixed agents fall in the same indifference class of a weak order. In Theorem 1, I show that the relative frequency $\tilde{T}_n/T_n$ approximates the probability that two men admit a Pareto-improvement cycle (chain) in a decentralized market.

Next I approach Pareto-improvement cycles and chains more rigorously by offering a formal definition based on the work of Erdil and Ergin (2015). To simplify the discussion, it is useful to first define the concept of an *envy relation*. A man m_1 *weakly envies* another man m_2 if m_1 weakly prefers m_2's partner to his own match, so that $\mu(m_2)R_{m_1}\mu(m_1)$. Envy relations are denoted with an arrow ($m_1 \rightarrow m_2$); strict envy is analogously defined with a strict preference.[6] The arrow notation makes Pareto-improvement cycles and chains easy to visualize using the following definition.

Definition 3. Given a matching μ, a **Pareto-improvement cycle** is a sequence of men $m_1, m_2, \ldots, m_K$ ($K \geq 2$) and their respective partners $\mu(m_1), \mu(m_2), \ldots, \mu(m_K)$, such that:

1. $\mu(m_1) \rightarrow \mu(m_2) \rightarrow \ldots \rightarrow \mu(m_K) \rightarrow \mu(m_1)$
2. $m_K \rightarrow m_{K-1} \rightarrow \ldots \rightarrow m_2 \rightarrow m_1 \rightarrow m_K$,

where at least one of the envy relations is strict.

A **Pareto-improvement chain** is a set of men $m_1, \ldots m_K$ and women $\mu(m_2)$, $\mu(m_3)$, …, $\mu(m_K)$, w_K, ($K \geq 2$) such that:

[6] For the case where there are single agents, a married agent can also point to (or "weakly envy") a single one. In this case, we say $m \rightarrow w$ if m weakly prefers the single woman w to his current match. Strict envy is defined analogously.

1. m_1 and w_K are single;
2. $m_1 \to m_2, \dots, m_{K-1} \to m_K$ and $m_K \to w_K$.
3. $w_K \to \mu(m_K) \to \mu(m_{K-1}) \to \dots \to \mu(m_1)$.
4. At least one of the envy relations is strict.

Pareto-improvement cycles and chains are therefore sets of agents who are willing to exchange partners amongst themselves so that at least one agent in the cycle (chain) is better off and nobody else is worse off. Using arrows to denote envy relations, Pareto-improvement cycles and chains can be represented more intuitively as graphs (Fig. 1 and 2).

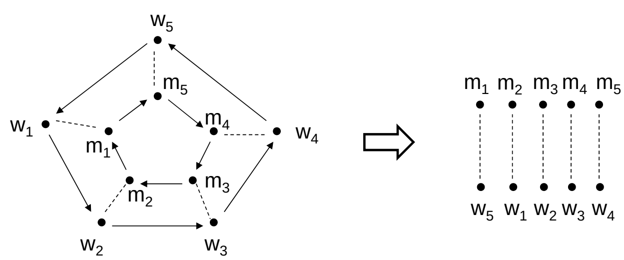

Figure 1: A Pareto-improvement cycle (left panel) and the Pareto-improving assignment it suggests (right panel).

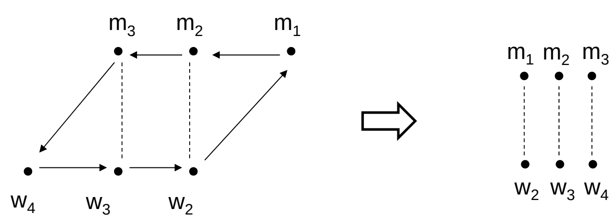

Figure 2: A Pareto-improvement chain (left panel) and the Pareto-improving assignment it suggests (right panel).

Whether the graph is a cycle or a chain is closely linked to the concept of *acceptability*. A man m is *acceptable* to a woman w if she ranks him strictly better than remaining single. When all agents are acceptable, chains cannot form, since no individual wishes to remain single. To facilitate the exposition, I will assume throughout that all agents are acceptable, and state the main theorems in terms of cycles, but they remain equally applicable to chains; the proof is relegated to the Appendix.

Pareto-improvement cycles and chains provide a means to construct a matching that Pareto dominates a given matching μ, so if a cycle or chain exists, it is immediate that the matching μ is inefficient. The converse is also true, so a key feature of Pareto-improvement cycles and chains is that they identify inefficient matchings:

Lemma 1 (Erdil and Ergin (2015), Theorem 1.) *A one-to-one matching is Pareto efficient if and only if it admits neither Pareto-improvement cycles nor Pareto-improvement chains.*

Proof. See Erdil and Ergin (2015), Theorem 1.

The size of the cycle or chain is not indicative of the number of agents who will strictly improve following the Pareto-improving exchange suggested by the cycle. For example, it may be the case that only one of the envy relations in the cycle in Fig. 1 (left) is strict, but nonetheless all 10 men and women involved may have to change partners for one particular agent to obtain a better mate. In most cases, therefore, the number of agents involved in Pareto-improvement cycles will overstate the true extent of the inefficiency. Likewise, it is also possible that most envy relations are strict, in which case most agents involved in the cycle will improve strictly. This problem is akin to some issues identified in the kidney exchange literature: sometimes a patient may not be able to get a compatible donor unless a very large chain of exchanges takes place. Since the population remaining outside Pareto-improvement cycles cannot improve from a reassignment, the fraction of agents who admit Pareto-improvement cycles (chains) is a suitably conservative measure of inefficiency.

Similarly, nothing prevents agents from simultaneously participating in different cycles or chains of various size. In this context, it would make sense to look for the smallest cycle or chain that can make a fixed number of agents strictly better off. To reflect this, I focus on the smallest building component of a cycle: two agents from the same side of the market along with their mates. Two men from a stable matching cannot be part of *any* cycle unless at least one of their mates is indifferent between them, because otherwise the matching is either unstable, or else their mates would be unwilling to switch.

Lemma 2 *(Necessary condition for cycle formation). Suppose that μ is a stable matching and men have strict preferences. Then a given 2-tuple of men (m_1, m_2) cannot participate in any Pareto-improvement cycle unless at least one of the partners $\mu(m_1)$ and $\mu(m_2)$ is indifferent between m_1 and m_2, so that*

$$m_1 \sim_{\mu(m_2)} m_2 \quad \text{or} \quad m_1 \sim_{\mu(m_1)} m_2, \quad \text{or both .}^7 \tag{1}$$

When this sufficient condition is met, we say that the 2-tuple (m_1, m_2) admits a cycle.

[7] This "non-exclusive OR" and the notation $\cup$ are used interchangeably in the proofs.

Proof: See the Appendix.

Clearly, admitting a cycle is only a minimum condition and does not guarantee that a cycle exists. Therefore, checking if (1) holds for every distinct 2-tuple of men is a suitably conservative way of gauging the Pareto efficiency of a stable match. It is convenient to further disaggregate condition (1) in two parts. Rather than checking each unordered pair (m_1, m_2) for the two indifferences in (1), it is simpler to consider *ordered* pairs of men instead, associating one ordered pair with each indifference, and check each ordered pair for *one* indifference per the following definition:

Definition 4. An ordered 2-tuple (m_1, m_2) is said to *admit a cycle* if $m_1 \sim_{\mu(m_2)} m_2$.

Clearly, if no ordered 2-tuple admits a cycle, no unordered 2-tuple admits a cycle either; and if no two agents in the problem admit a cycle, then no cycle (of any length) exists, and the matching is Pareto efficient.

Lemma 2 provides a partial link between cycles and indifferences in agent preferences, but the link is not complete, because Lemma 2 contains no information about the assignment of partners under μ. The condition $m_1 \sim_{\mu(m_2)} m_2$ could be read in n different ways, depending on whether m_2 is matched to $w_1, w_2, \ldots,$ or w_n. To find how often two given men will admit a cycle, one needs to know how frequently the random stable mechanism assigns each woman to each man — both unconditionally and conditional on a pairwise indifference. The next two results answer this question.

Lemma 3 ***(Symmetry lemma)***. *Let agent preferences be selected according to the preference-generating process $\mathscr{P}_0$. Then the random stable mechanism assigns any given woman to any given man with equal probability, so that*

$$\Pr\big(\mu(w) = m\big) = \frac{1}{n} \quad \forall (m, w). \tag{2}$$

Proof: See the Appendix.

Lemma 3 shows that imposing a stability filter on a random matching process preserves its symmetric properties. However, this statement is unconditional. If one were to condition a woman w's weak preferences — for example, by restricting her choice only among those orders where m_1 and m_2 are equivalent — then symmetry will not hold, because some marriages will be made more likely. For instance, any matching in which w is assigned to

m_2 will *never* be blocked by the pair (m_1, w) because of the above indifference; whereas if w's preferences were unrestricted, the same assignment would be blocked with a positive probability, therefore making its stability and consequent selection by the random stable mechanism less likely.

Since the same outcome — a particular man being assigned to a particular woman — occurs with a different probability depending on whether or not it is conditioned on an indifference, there is clearly a statistical dependence between the allocation of men to women and pairwise indifferences in women's preferences. This renders Lemma 2 of limited usefulness, as the frequency of a cycle cannot be immediately related to a known quantity.

Fortunately, in large markets with independent preferences, there is sufficient heterogeneity for this statistical dependence to become vanishing, which allows one to disentangle the influence of preferences on the allocation from the influence of the mechanism. In turn, this allows one to approximate the probability of two agents admitting a cycle with the probability of a pairwise indifference, which depends only on the preference-generating process.

Theorem 1 *a) The stochastic allocation μ_n attained by the random stable mechanism in a marriage problem of size n is asymptotically independent of the occurrence of indifferences in the underlying preference profile in the sense that, as $n \to \infty$, for every two men m_1 and m_2 and every woman w, it is true that*

$$\Pr\left(\mu_n(m_2) = w \cap m_1 \sim_w m_2\right) \to \Pr\left(\mu_n(m_2) = w\right) \Pr(m_1 \sim_w m_2). \tag{3}$$

b) This implies that, for any two arbitrary men m_1, m_2 and arbitrary woman w, as $n \to \infty$,

$$\Pr\left(m_1 \sim_{\mu_n(m_2)} m_2\right) \to \Pr\left(m_1 \sim_w m_2\right). \tag{4}$$

Proof. See the Appendix.

The intuition behind the result in part (a) is that, even though the pair (m_1, w) never blocks the marriage between m_2 and w when $m_1 \sim_w m_2$, nonetheless, as the number of agents grows, chances increase that at least one other pair, whose preferences are unrestricted, will block the same marriage. Thus, the marginal absence of a single blocking pair becomes negligible for large n, and the allocation of men to women is almost symmetric. Since the mechanism's outcome retains its symmetry regardless of whether it is conditioned on a fixed pairwise indifference or not, this also implies that the allocation is independent of pairwise indifferences — a fact that permits a more direct link between the frequency of cycles and that of indifferences.

Theorem 1(b) also shows that, asymptotically, the necessary condition for cycles in Lemma 2 (which depends on the matching μ) and the indifference draw $m_1 \sim m_2$ by an arbitrary woman (which does not depend on the matching), are almost the same thing. This permits a direct link between cycles and indifferences in the underlying preference, which can be computed based on the (known) preference-generating process.

To make use of Theorem 1, one needs to compute the probability $\Pr(m_1 \sim_w m_2)$ of indifference between two arbitrary men by an arbitrary woman w. When preferences are uniform, *i.i.d.*, this probability equals

$$\begin{aligned}\Pr(m_1 \sim_w m_2 \ ; \ n) &= \frac{\#\text{ of weak orders where } m_1 \text{ and } m_2 \text{ are in the same partition block}}{\text{total } \#\text{ of weak orders}} \\ &\equiv \frac{\tilde{T}_n}{T_n},\end{aligned} \tag{5}$$

where $\tilde{T}_n$ denotes the number of weak preferences where $m_1 \sim m_2$ and T_n is the total number of weak preferences (orders).

For example, for a market of size 15,

$$\Pr(m_1 \sim_w m_2 \ ; \ n = 15) = \frac{10,641,342,970,443}{230,283,190,977,853} = 0.0462. \tag{6}$$

Two things stand out in this example. First, the number of possible weak preferences over only 15 agents is already around 230 trillion (about 176 times larger than the number of linear orders, 15!) because of the very large number of ways that agents can be grouped together in indifference classes and those classes arranged in order of preference. This creates a significant computational challenge, so a tractable way of finding the numbers T_n and $\tilde{T}_n$ is essential. Secondly, despite the large values of $\tilde{T}_n$ and T_n, their ratio is relatively small: just below 5%. This is sufficiently low to suggest that perhaps asymptotically, the probability of a cycle will converge to zero. The next theorem provides a technology for counting the total number of ordered partitions T_n and the number of cases $\tilde{T}_n$ where two fixed men occur in the same block of the ordered partition. As it turns out, there is no closed formula for either T_n or $\tilde{T}_n$, but they still can be found using recurrence relations. Moreover, it turns out that both numbers can be expressed in terms of the same numerical sequence.

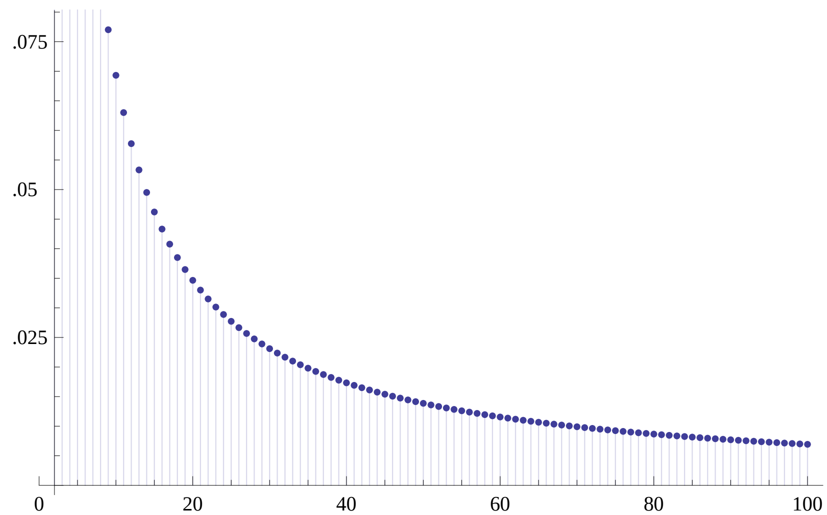

Figure 3: The sequence T_{n-1}/T_n. (Source: author's computations.)

Theorem 2 *(a)The total number T_n of weak preferences over n partners satisfies the recurrence relation*

$$T_n = \sum_{i=0}^{n-1} \binom{n}{i} T_i. \tag{7}$$

(b) The total number $\tilde{T}_n$ of weak preferences over n partners in which two fixed agents are in the same indifference class satisfies the recurrence relation

$$\tilde{T}_n = T_{n-1}. \tag{8}$$

Proof: See the Appendix.

The second part of the theorem allows one to reformulate the probability in terms of a single numerical sequence, T_n. It implies that

$$\Pr(m_1 \sim_w m_2 \ ; \ n) = \frac{T_{n-1}}{T_n}, \tag{9}$$

thereby making the probability of a pairwise indifference equal to the ratio of two successive terms in the sequence. The question therefore becomes whether T_n grows sufficiently fast to make the ratio of two successive terms vanishing. One can form a conjecture about the asymptotic behavior of this probability by computing and plotting its first few terms numerically, as shown in Fig. 3. The figure suggests that the sequence likely converges to zero; that this is indeed the case is verified by the next theorem.

Theorem 3 *Under the default preference-generating process $\mathscr{P}_0$, the probability that an arbitrary woman w is indifferent between two given men m_1 and m_2 satisfies the limit*

$$\lim_{n\to\infty} \Pr(m_1 \sim_w m_2\ ;\ n) = \lim_{n\to\infty} \frac{T_{n-1}}{T_n} = 0. \tag{10}$$

Proof: See the Appendix.

Corollary 1 *Theorems 1 and 3 imply that the probability that an arbitrary ordered 2-tuple of men (m_1, m_2) admits a Pareto-improvement cycle also converges to zero:*

$$\lim_{n\to\infty} \Pr(m_1 \sim_{\mu(w_2)} m_2\ ;\ n) = 0. \tag{11}$$

According to the corollary, the chance that an arbitrary ordered 2-tuple permits a cycle becomes vanishing for large n. Now it remains only to relate the result obtained for 2-tuples to individual agents, which provide a more meaningful metric for gauging the inefficiency. I am particularly interested in two inefficiency metrics: the absolute number and the proportion of men who admit Pareto-improvement cycles. Since, for each market size n, the final allocation μ is stochastic (on account of the random preferences and the randomness of the mechanism), I look at these metrics in expectation.

Theorem 4 *Let α_n denote the expected number of men involved in ordered 2-tuples that admit cycles in a problem of size n, and let agent preferences be drawn according to the default preference-generating process $\mathscr{P}_0$. Then, as $n \to \infty$:*

a) The expected number of men α_n who admit cycles increases with n:

$$\alpha_n = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}} \ \to\ \infty. \tag{12}$$

b) The expected proportion of men $\frac{\alpha_n}{n}$ who admit cycles converges to 0:

$$\frac{\alpha_n}{n} = \frac{1}{2n} + \frac{1}{2n}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}} \ \to\ 0. \tag{13}$$

Proof: See the Appendix.

The first striking feature of this result is that the expected number of men who admit cycles grows without bound, despite the vanishing probability of a cycle. The reason for this is the slow rate of convergence of pairwise indifferences to zero, compared with the speed at which ordered 2-tuples grow with n. Since ordered 2-tuples grow of the order of n^2, while the quantity T_{n-1}/T_n goes to zero of the order of $1/n$, the expected number of cycle-admitting 2-tuples increases too quickly with n. As a result, both the number of cycle-admitting 2-tuples and the number of men involved in them grow without bound. This might mislead one to think that all inefficiency metrics behave the same way and that the matching is inefficient; however, this is not the case.

As shown in part (b), pairwise indifferences still vanish quickly enough to make the *proportion* of cycle-admitting men vanish. Given that T_{n-1}/T_n is of the order of $1/n$, the number of cycle-admitting men α_n according to equation (12) grows of the order of $\sqrt{n-1}$, which is not fast enough to survive another division by n. Therefore the expected proportion α_n/n of cycle-admitting men goes to zero and the matching becomes asymptotically Pareto efficient, albeit at a slow rate. This rate of convergence and its practical implications are discussed in detail in section 4.

This underscores the non-triviality of the main result and disproves the intuitive fallacy that the zero limit of pairwise indifferences predetermines the outcome. The *rate of convergence* to zero — a feature determined by the combinatorial properties of ordered partitions — is much more important.

This discussion naturally opens the question of what happens with preferences where the chance of a pairwise indifference converges to zero faster than $1/n$. Preferences with a larger (smaller) number of ties can arise naturally in the context of decision costs, where it is more (less) costly to perform pairwise comparisons, dictating the use of fewer or more "bins."[8] As shown in the next theorem, if the sequence T_{n-1}/T_n was instead of the order of $1/n^2$, the number of cycle-admitters would converge to a positive constant; and if T_{n-1}/T_n converged of the order of $1/n^3$ or faster, then both the number of men and the fraction of men engaged in Pareto-improving cycles vanish for large n. However, not all preference-generating processes with faster convergence are symmetric, whereas symmetry is essential for making the analysis tractable, as evident from Theorem 1. Therefore, it is still necessary to place some restrictions on preferences.

Definition 5. A stochastic PGP is **symmetric** if and only if $\Pr(aPb) = \Pr(cPd)$ and $\Pr(a \sim b) = \Pr(c \sim d)$ for any quadruple of (not necessarily distinct) agents (a, b, c, d).

[8]Employee recruiting and selection is one practical example.

Theorem 5 *If a stochastic PGP for weak preferences $\mathscr{P}^*$ is symmetric and agent preferences are mutually independent, then the following statements hold:*

a) If the probability of a pairwise indifference $\Pr(m_1 \sim_w m_2)$ *under $\mathscr{P}^*$ is of the order of $1/n^2$, then $\alpha_n \to const.$ and $\frac{\alpha_n}{n} \to 0$ as $n \to \infty$.*

b) If the probability of a pairwise indifference $\Pr(m_1 \sim_w m_2)$ *under $\mathscr{P}^*$ is of the order of $1/n^k$ for $k \geq 3$, then $\alpha_n \to 0$ and $\frac{\alpha_n}{n} \to 0$ as $n \to \infty$.*

Proof: See the Appendix.

An example of a non-uniform but symmetric weak-preference PGP where the chance of a pairwise indifference converges to zero at a rate faster than $1/n$ can be devised as follows. Split the list of all T_n weak orders into two groups: Group 1 containing only strict (linear) orders, and Group 2 comprising all orders containing at least one indifference. When selecting a preference, first select a group at random, where Group 1 is chosen with probability $1 - \frac{1}{n}$, and Group 2 with probability $\frac{1}{n}$. Next, having chosen a group, select a preference uniformly at random from that group.

With this PGP, it is easy to show that the probability of a pairwise indifference within Group 2 is

$$\frac{T_{n-1}}{T_n - n!} \to \frac{T_{n-1}}{T_n}, \tag{14}$$

so the chance of a pairwise indifference within Group 2 still goes to zero of the order of $1/n$. However, since Group 2 itself is selected with probability $1/n$, the overall chance of a pairwise indifference is now $\frac{1}{n}\frac{T_{n-1}}{T_n}$, which is of the order of $1/n^2$.

By weighting Group 1 and Group 2 with respective probability weights of $1 - \frac{1}{n^k}$ and $\frac{1}{n^k}$, one can similarly generate preferences where the chance of a pairwise indifference goes to zero at any desired rate of $1/k$; the process is symmetric by construction, since the orders contained within each group are fully symmetric.

4 Discussion of Results

Theorems 4 and 5 permit a practical estimate of how large a market is large enough to reduce the inefficiency to acceptable levels. Table 1 shows the expected proportion of cycle-admitting men in the benchmark case with uniform preferences. According to the table,

Proportion of cycle-admitting men as a function of market size n									
n	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90
	-	0.2891	0.2027	0.1641	0.1412	0.1257	0.1142	0.1054	0.0983
	-	0.2755	0.1976	0.1613	0.1394	0.1244	0.1133	0.1046	0.0977
	0.5928	0.2636	0.1929	0.1587	0.1376	0.1231	0.1123	0.1039	0.0971
	0.5066	0.2531	0.1885	0.1561	0.1360	0.1219	0.1114	0.1031	0.0964
	0.4477	0.2437	0.1844	0.1537	0.1343	0.1207	0.1104	0.1024	0.0958
	0.4046	0.2352	0.1805	0.1514	0.1328	0.1195	0.1096	0.1017	0.0953
	0.3714	0.2275	0.1768	0.1492	0.1313	0.1184	0.1087	0.1010	0.0947
	0.3448	0.2205	0.1734	0.1471	0.1298	0.1173	0.1078	0.1003	0.0941
	0.3230	0.2141	0.1701	0.1450	0.1284	0.1163	0.1070	0.0996	0.0936
	0.3047	0.2082	0.1670	0.1431	0.1270	0.1152	0.1062	0.0990	0.0930

Table 1: The maximum expected proportion of men who admit Pareto-improvement cycles, as provided by the sequence α_n/n in equation (13). (Source: author's computations.)

the fraction of cycle-admitters falls quickly before flattening out, and for n larger than 79 is already below 10%.[9] However, even this number in fact overstates the true inefficiency.

To see this, first observe that not every two men who *admit* a cycle actually *participate* in one. Admitting a cycle means simply that the necessary condition for cycles (from Lemma 2) has been met; this in itself does not imply that a cycle is present. Secondly, even when a cycle is present, not everyone in the cycle needs to improve strictly, since only one of the envy relations needs to be strict. What is required to form a cycle is that only a single agent improve while no one else is made worse off, which can result in potentially very large cycles with only a single agent who benefits strictly from the exchanges.[10] Therefore, the numbers from Table 1 correspond to an absolute worst-case scenario. This suggests that even in moderate-sized markets with weak preferences, Pareto inefficiency is not a major concern as long as preferences remain uncorrelated.

Naturally, my approach also has some limitations. From the discussion of Theorem 1 it is clear that anything that upsets the symmetry of the mechanism or the symmetry and statistical independence of individual preferences also has the potential to affect the main result.[11] Despite that, I consider the random stable mechanism with a large class of non-

[9]For the case where both sides of the market have weak preferences, the numbers in the table are doubled, as shown in Theorem 6 in the Appendix.

[10]This aspect is very similar to the kidney exchange literature: sometimes a patient may not be able to obtain a compatible donor unless a very large chain of exchanges takes place.

[11]Other than in very special cases where asymmetric preferences and asymmetries in the selection of a stable match cancel each other out.

uniform random preferences and show that, as long as the preference-generating process is symmetric and preferences remain independent, the same or stronger results hold.

I do not focus on other random mechanisms (e.g., the men-optimal mechanism with tie-breaking) because the distribution of matched pairs in such multi-step algorithms is *path-dependent*. Since a girl can only compare offers from current proposers, the final allocation depends not only on the presence of ties in each woman's preference relation, but also on how such ties were resolved in previous steps. For example, suppose a tie between two men m' and m'' is resolved by a woman w' in favor of m' at step $k-1$ of the men-optimal algorithm. This implies that m'' gets to propose to his next-best choice (say w'') at the next step k. Suppose also that w'' happens to be indifferent between m'' and some other proposer m'''; but if the first tie was resolved the other way and m'' ended up proposing to w'' a few steps later (when a strictly better proposer arrives and bumps him down), then m'' could be part of an entirely different proposer set, where the same woman w'' does not have any indifferences. Since the technology in this paper does not account for path-dependence, it is unsuitable for analyzing this question.

Unlike most results related to the marriage problem, my result also does not automatically extend to many-to-one matching. To see this, consider an example with n students with weak preferences applying to m colleges, where $n \geq m$. A necessary condition for two colleges c_1 and c_2 to admit a cycle is that at least one student from the set S_1 of students admitted to c_1 must be indifferent between schools c_1 and c_2. The relative frequency of this event is bounded by the probability

$$\Pr\left(\bigcup_{i\in S_1} c_1 \sim_i c_2\right) \leq \sum_{i\in S_1} \Pr(c_1 \sim_i c_2) = |S_1| \Pr(c_1 \sim_i c_2).$$

Whether the sum on the right-hand side converges to zero depends on whether the cardinality of the set S_1 changes faster than the probability $\Pr(c_1 \sim_i c_2)$. For example, assume each college has the same fixed capacity $\bar{c}$ and students increase in steps of $\bar{c}$, so that a new college is added per every $\bar{c}$ additional students; then $|S_1|$ grows of the same order as n, while the probability $\Pr(c_1 \sim_i c_2)$ is of the order of $1/n$, causing the right-hand side to converge to a positive constant that renders the upper bound uninformative.[12] Even if colleges do not all have equal capacities, the average capacity $\mathbb{E}|S_1|$ will still grow of the order of n, so there will exist schools for which the probability of a cycle does not converge to zero.

[12] If instead we choose to have a fixed number of colleges with increasing capacities, then $|S_1|$ will still grow of the order of n and the right-hand side will again tend to a non-zero constant.

These limitations help outline more clearly the role of the paper in the matching literature. Firstly, it demonstrates that the absence of a central matchmaker is not necessarily bad news for the Pareto efficiency of the random outcome in a decentralized weak-preference market. The analysis shows that even moderate-sized markets with size as low as 79 in practice attain efficiency levels above 90%, as measured by the fraction of agents who cannot Pareto-improve from a reassignment. This is a stark result, keeping in mind that the methodology provides a loose lower bound on efficiency, corresponding to an absolute worst-case scenario. Similar to Roth and Vande Vate (1990), who show that the random nature of meetings in a decentralized market is not an obstacle to achieving stability, this paper demonstrates that randomness is analogously helpful in achieving (approximately) Pareto efficient outcomes under weak preferences.

5 Conclusion

In decentralized markets without a coordinating body, stability can create perverse incentives for inefficiently matched agents to stay together. This implies that stable, but inefficient matches have no tendency to dissolve or improve their efficiency over time. Unlike random paths to stability, there is no obvious process by which an inefficient stable match could evolve into a more Pareto efficient one. This raises the question of whether a decentralized market with weak preferences can attain Pareto efficiency in the absence of a central matchmaker.

To answer this question, I approach it similarly to Kojima and Pathak (2009) and Kojima and Manea (2010), who show that some shortcomings of stable two-sided matchings, such as the lack of strategy-proofness, tend to improve in large markets. In the spirit of this literature, I demonstrate that with weak preferences, a uniformly selected random stable matching will be approximately Pareto efficient in a large decentralized market.

In practice, even moderate-sized markets can attain good efficiency levels. The market size necessary to reduce the proportion of agents admitting Pareto-improving exchanges below 10% is only $n = 79$ when one side of the market admits weak preferences, and $n = 158$ when both sides admit weak preferences. This implies that approximate Pareto efficiency is attainable in a decentralized market even in the absence of a central matchmaker.

6 Appendix

6.1 Theorem Proofs

Lemma 2 *(Necessary condition for cycle formation.) Suppose that μ is a stable matching and men have strict preferences. Then a given 2-tuple of men (m_1, m_2) cannot participate in any Pareto-improvement cycle unless at least one of the partners $\mu(m_1)$ and $\mu(m_2)$ is indifferent between m_1 and m_2, so that*

$$m_1 \sim_{\mu(m_2)} m_2 \quad \text{or} \quad m_1 \sim_{\mu(m_1)} m_2, \text{ or both.}$$

Proof. First consider the preferences of m_1. Given any matching μ, the preference is either $\mu(m_2)P_{m_1}\mu(m_1)$, or $\mu(m_1)P_{m_1}\mu(m_2)$, because men's preferences are strict (the case where both men and women admit weak preferences is addressed in Theorem 5).

Initially, suppose that $\mu(m_1)P_{m_1}\mu(m_2)$; then m_1 will not be willing to trade with m_2, and m_1 and m_2 do not admit a Pareto-improvement cycle.

Now suppose instead that $\mu(m_2)P_{m_1}\mu(m_1)$ and consider the weak preferences of $\mu(m_2)$. If $m_1 P_{\mu(m_2)} m_2$, then the pair $(m_1, \mu(m_2))$ is a blocking pair and μ is not stable, a contradiction. If $m_2 P_{\mu(m_2)} m_1$, then $\mu(m_2)$ cannot point to $\mu(m_1)$, so by Definition 3, m_1 and m_2 do not admit a Pareto-improvement cycle. The only remaining preference of $\mu(m_2)$ consistent with a Pareto-improvement cycle is $m_1 \sim_{\mu(m_2)} m_2$. Reversing the places of m_1 and m_2 yields the second part of the necessary condition, $m_1 \sim_{\mu(m_1)} m_2$. □

Lemma 3 *Let agent preferences be selected according to the preference-generating process $\mathscr{P}_0$. Then the random stable mechanism assigns any given woman to any given man with equal probability, so that*

$$\Pr\big(\mu(w) = m\big) = \frac{1}{n} \quad \forall (m, w). \tag{15}$$

Proof. The proof is by contradiction. Suppose the random stable mechanism is not symmetric. Then there exists a woman w^* and men m^* and m'^* such that w^* is assigned to m^* more frequently than to m'^* in a stable matching, so that

$$\Pr\big(\mu(w^*) = m^*\big) > \Pr\big(\nu(w^*) = m'^*\big), \tag{16}$$

where μ and ν denote the respective different assignments produced by the random stable mechanism.

Since the realized stable matching μ is chosen from the uniform distribution over all stable matchings (as determined by each joint preference draw), each stable matching has an equal chance of being picked. Then the asymmetry (16) means that the assignment $\nu(w^*) = m'^*$ is blocked more frequently than the assignment $\mu(w^*) = m^*$. This implies that there exists at least one pair, denoted without loss of generality as (m, w), that blocks the assignment ν of woman w^* to m'^* more frequently than her assignment μ to man m^*.

For the man and woman (m, w) to block the fixed assignment ν (and all matchings containing it), they must prefer each other strictly, relative to the respective partners assigned to them under ν. Hence for the ν to be blocked more frequently by the pair (m, w), it must be true that

$$\Pr\left(\bigcap\begin{matrix} w\ P_m\ \nu(m) \\ m\ P_w\ \nu(w)\end{matrix}\right) > \Pr\left(\bigcap\begin{matrix} w\ P_m\ \mu(m) \\ m\ P_w\ \mu(w)\end{matrix}\right). \tag{17}$$

Combining (17) with the fact that preferences P_m and P_w are drawn independently implies that

$$\begin{aligned}\Pr\left(\bigcap\begin{matrix} w\ P_m\ \nu(m) \\ m\ P_w\ \nu(w)\end{matrix}\right) &= \Pr\big(w\ P_m\ \nu(m)\big)\Pr(m\ P_w\ \nu(w)\big) > \\ > \Pr\left(\bigcap\begin{matrix} w\ P_m\ \mu(m) \\ m\ P_w\ \mu(w)\end{matrix}\right) &= \Pr\big(w\ P_m\ \mu(m)\big)\Pr\big(m\ P_w\ \mu(w)\big)\end{aligned} \tag{18}$$

or in other words, that

$$\Pr\big(w\ P_m\ \nu(m)\big)\Pr\big(m\ P_w\ \nu(w)\big) > \Pr\big(w\ P_m\ \mu(m)\big)\Pr\big(m\ P_w\ \mu(w)\big). \tag{19}$$

Each of these probabilities expresses the chance of one fixed element preceding another in the respective preferences of the man m and woman w. However, since every possible preference is equally likely, preferences are symmetric and therefore[13]

$$\begin{aligned}&\Pr\big(w\ P_m\ \nu(m)\big) = \Pr\big(w\ P_m\ \mu(m)\big) \text{ and} \\ &\Pr\big(m\ P_w\ \nu(w)\big) = \Pr\big(m\ P_w\ \mu(w)\big),\end{aligned} \tag{20}$$

which implies that (19) should hold with equality, a contradiction.

Therefore a random stable matching must be symmetric in that each man and each woman are matched with the same probability as any other man-woman pair. □

[13] To prove equation (20) formally, let the number of strict (weak) orderings of the set of women where any given woman w precedes another w', be denoted by N_s (N_w for the weak order, respectively). Since the two elements w and w' are arbitrary, we can replace the pair w, w' with the pair w, w'' and the total number of cases where w precedes w'' would still be N_s (N_w, respectively). Hence, there are just as many strict (weak) orderings where w precedes w' as there are strict (weak) orderings where w precedes w'', and because preferences are picked uniformly, $\Pr(wPw') = \Pr(wPw'')$.

Theorem 1

a) *The stochastic allocation μ_n attained by the random stable mechanism in a marriage problem of size n is asymptotically independent of the occurrence of indifferences in the underlying preference profile in the sense that, as $n \to \infty$, for every two men m_1 and m_2 and every woman w, it is true that*

$$\Pr\big(\mu_n(m_2) = w \cap m_1 \sim_w m_2\big) \to \Pr\big(\mu_n(m_2) = w\big)\Pr(m_1 \sim_w m_2). \tag{21}$$

b) *This implies that, for any two arbitrary men m_1, m_2 and arbitrary woman w, as $n \to \infty$,*

$$\Pr\big(m_1 \sim_{\mu_n(m_2)} m_2\big) \to \Pr\big(m_1 \sim_w m_2\big). \tag{22}$$

Proof. (a). I construct the proof by showing that the probability that a matching containing the fixed assignment $\nu_n(w) = m_2$ is stable does not depend on whether one conditions on a particular fixed indifference or not.

Let $\{\nu_n\}$ be a sequence of matchings of increasing size n, containing the fixed assignment $\nu_n(w) = m_2$. For each problem size n, I fix the matching ν_n and let agent preferences vary. For ν_n to be stable, no blocking pairs should exist. For a problem of size n, there are $n(n-1)$ potential blocking pairs, since a man will never form a blocking pair with his wife. The potential blocking pairs and married pairs could be visualized in the set $M \times W$ as follows:

$$\begin{array}{cccc} (m_1, w_1) & \boldsymbol{(m_1, w_2)} & \cdots & (m_1, w_n) \\ \boldsymbol{(m_2, w_1)} & (m_2, w_2) & \cdots & (m_2, w_n) \\ \vdots & \vdots & \ddots & \vdots \\ (m_n, w_1) & (m_n, w_2) & \cdots & \boldsymbol{(m_n, w_n)} \end{array} \tag{23}$$

where the bolded pairs are married pairs from an example matching

$$\eta(m_1, m_2, \ldots, m_n) = (w_2, w_1, \ldots, w_n),$$

and the remaining $n(n-1)$ pairs are potential blocking pairs that can block the example matching if they draw the right combination of preferences. For any fixed matching, we group the matched pairs in a set $\mathcal{V} \subset M \times W$. The probability that a given matching ν_n is stable under the standard PGP with unrestricted preferences $\mathscr{P}_0$ can therefore be expressed as

$$\Pr(\nu_n \text{ is stable } |\mathscr{P}_0) = \Pr\left(\bigcap_{(m,w)\in(M\times W)\setminus\mathcal{V}} (m, w) \ \text{ does not block } \ \nu_n\right). \tag{24}$$

I will show that this probability has the same limit, regardless of whether it is conditioned a fixed pairwise indifference (which is required to enable cycles). To condition on such an indifference, I define a restricted PGP $\mathscr{P}_1$.

DEFINITION 6. The restricted preference-generating process $\mathscr{P}_1$ is defined as follows:

a) For every $m \in M$, $\mathscr{P}_1^m$ is the same as $\mathscr{P}_0^m$.

b) For every woman $\omega \neq w$, $\mathscr{P}^\omega = \mathscr{P}_0^\omega$. For woman $\omega = w$, the preference R_w is drawn from the uniform distribution over the subset of weak preferences where $m_1 \sim_w m_2$.

c) Individual agents' draws are mutually independent.

Showing that the probability $\Pr(\nu_n \text{ is stable } | \mathscr{P}_0) \to \Pr(\nu_n \text{ is stable } | \mathscr{P}_1)$ will imply that the stability of ν_n is independent of the occurrence of a fixed pairwise indifference when n is large. This is the crux of the proof.

However, the probability (24) cannot be easily decomposed as a product, because potential blocking pairs are not all independent in blocking ν_n. The reason for this is that some pairs contain overlapping agents, whose preferences introduce correlation. For example, pairs (m_1, w_1) and (m_1, w_3) are not independent in blocking the example matching η because the man's necessary condition to prefer w_1 to his current match also depends on whether he likes w_3 better than his current match:

$$\Pr\left(w_1 P_{m_1} \eta(m_1) \mid w_3 P_{m_1} \eta(m_1)\right) = \frac{2}{3} \neq \frac{1}{2} = \Pr\left(w_1 P_{m_1} \eta(m_1)\right). \tag{25}$$

On the other hand, given a matching ν_n, I claim that there are always at least n potential blocking pairs in the set $(M \times W) \backslash \mathcal{V}$ that are mutually independent in blocking ν_n, as these pairs share no agents in common. One can always select a set of n such pairs according to the following algorithm:

1. For every matched pair $(m_k, w_j) \in \mathcal{V}$, select its right-hand adjacent pair (m_k, w_{j+1}) if $j + 1 \leq n$;

2. Where $j + 1 > n$, select the leftmost pair in the same row, i.e., (m_k, w_1).

This algorithm ensures that, from each row and each column of the matrix $(M \times W) \backslash \mathcal{V}$, exactly one potential blocking pair is selected, which ensures no agents in common and guarantees mutual independence of agent preferences. Therefore, we will split the set of potential blocking pairs, $(M \times W)\backslash\mathcal{V}$, into a subset of mutually independent pairs $\mathcal{I}$ and all remaining pairs $\mathcal{R}$. Let us denote with A_n the event that ν_n is stable under the standard (unrestricted) uniform preference-selection process $\mathscr{P}_0$, and decompose A_n into an "independent" component $A_{I,n}$ and a remaining component $A_{R,n}$:

$$
\begin{aligned}
A_n &\equiv \left(\bigcap_{(m,w)\in(M\times W)\backslash\mathcal{V}} (m,w) \;\; \text{does not block} \;\; \nu_n \right) \\
A_{I,n} &\equiv \left(\bigcap_{(m,w)\in\mathcal{I}} (m,w) \;\; \text{does not block} \;\; \nu_n \right) \\
A_{R,n} &\equiv \left(\bigcap_{(m,w)\in\mathcal{R}} (m,w) \;\; \text{does not block} \;\; \nu_n \right).
\end{aligned} \tag{26}
$$

Then we can represent the event A_n that ν_n is stable as follows:

$$A_n = (A_{I,n} \cap A_{R,n}) \;\Rightarrow\; A_n \subset A_{I,n} \;\Rightarrow\; \Pr(A_n) < \Pr(A_{I,n}). \tag{27}$$

I will show that $\Pr(A_{I,n}) \to 0$ and therefore also $\Pr(A_n) \to 0$. Since the pairs in $\mathcal{I}$ are mutually independent in blocking ν_n, and the set $\mathcal{I}$ was constructed with cardinality n,

$$
\begin{aligned}
\Pr(A_{I,n}) \;&=\; \prod_{(m,w)\in\mathcal{I}} \Pr((m,w) \;\; \text{does not block} \;\; \nu_n) = \\
&= [\Pr\left((m,w) \;\; \text{does not block} \;\; \nu_n\right)]^n = \\
&= \left[1 - \Pr\left(\bigcap \begin{matrix} w \; P_m \; \nu_n(m) \\ m \; P_w \; \nu_n(w) \end{matrix}\right)\right]^n \to \\
&\to \left[1 - \frac{1}{2}\cdot\frac{1}{2}\right]^n = \left[\frac{3}{4}\right]^n \to 0,
\end{aligned} \tag{28}
$$

where in the last line we used the fact that for women, whose preferences are weak, the probability of one fixed mate being preferred to another tends to 1/2:

$$
\begin{aligned}
&\Pr(mP_w\nu_n(w)) + \Pr(\nu_n(w)P_w m) + \underbrace{\Pr(\nu_n(w) \sim_w m)}_{\to 0 \text{ by Theorem 3}} = 1 \Rightarrow \\
&\Rightarrow 2\Pr(mP_w\nu_n(w)) \to 1 \Rightarrow \Pr(mP_w\nu_n(w)) \to \frac{1}{2} \;\; \text{as } n \to \infty.
\end{aligned} \tag{29}
$$

Since $\Pr(A_{I,n}) > \Pr(A_n)$ and $\Pr(A_{I,n}) \to 0$, this implies

$$\Pr(A_n) \equiv \Pr(\nu_n \text{ is stable } | \mathscr{P}_0) \to 0. \tag{30}$$

Next I will show that the chance that ν_n is stable under the restricted preference-generating process $\mathscr{P}_1$ converges to the same limit, i.e.,

$$\Pr(\nu_n \text{ is stable } |\mathscr{P}_1) \to 0.$$

Consider the restricted preferences case ($\mathscr{P}_1$) where $m_1 \sim_w m_2$ and choice among preferences is uniform. Given the sequence of matchings ν_n, and recalling that it is specially constructed to always contain the fixed assignment $\nu_n(w) = m_2$, if we further restrict preferences to those orders where $m_1 \sim_w m_2$, then there is exactly one pair — the pair (m_1, w) — that will never block ν_n, all else equal, so that $\Pr\big((m_1, w) \text{ does not block } \nu_n\big) = 1$. Then we proceed as follows:

- If the pair $(m_1, w) \in \mathcal{I}$, then the exponent n in (28) becomes $n-1$, because one of the potential blocking pairs never blocks ν_n. This change does not alter the limit (28) as $n \to \infty$.
- If $(m_1, w) \in \mathcal{R}$, the proof in (28) holds without change, because the cardinality of $\mathcal{I}$ that drives the convergence is unaffected.

Thus we arrive at

$$\Pr(\nu_n \text{ is stable } |\mathscr{P}_1) \to 0. \tag{31}$$

Now let ε be a positive constant. Since under both $\mathscr{P}_0$ and $\mathscr{P}_1$, the probability that ν_n is stable has the same limit, given $\varepsilon > 0$, one can always find N large enough so that for $n \geq N$,

$$|\Pr(\nu_n \text{ is stable}|\mathscr{P}_1) - \Pr(\nu_n \text{ is stable}|\mathscr{P}_0)| < \varepsilon, \tag{32}$$

which implies that

$$\Pr(\nu_n \text{ is stable}|\mathscr{P}_1) \to \Pr(\nu_n \text{ is stable}|\mathscr{P}_0). \tag{33}$$

Now, given the key fact that the random stable mechanism selects among stable matchings uniformly at random, this means that

$$\begin{aligned} &\Pr(\nu_n \text{ is the selected stable matching by the RSM } |\mathscr{P}_1) \to \\ &\to \Pr(\nu_n \text{ is the selected stable matching by the RSM } |\mathscr{P}_0). \end{aligned} \tag{34}$$

However, recalling that the sequence $\{\nu_n\}$ is specially constructed so that each ν_n contains the fixed assignment $\nu_n(w) = m_2$, this implies that, as $n \to \infty$, the probability that the final matching μ_n selected by the random stable mechanism assigns w to m_2 converges to

$$\Pr\big(\mu_n(w) = m_2|\mathscr{P}_1\big) \to \Pr\big(\mu_n(w) = m_2|\mathscr{P}_0\big), \tag{35}$$

where $\mathscr{P}_1$ is the restricted PGP conditioned on a fixed indifference by w between m_1 and m_2, and $\mathscr{P}_0$ is the unrestricted preference process.

Written out fully, this means that, as $n \to \infty$,

$$\Pr\left(\mu_n(m_2) = w | m_1 \sim_w m_2\right) \to \Pr\left(\mu_n(m_2) = w\right), \text{ or} \tag{36}$$

$$\Pr\left(\mu_n(m_2) = w \cap m_1 \sim_w m_2\right) \to \Pr\left(\mu_n(m_2) = w\right) \Pr(m_1 \sim_w m_2), \tag{37}$$

which is the asymptotic independence property we were seeking to prove.

(b) Recall we already proved that the RSM is symmetric, so that for every pair (m, w), it is true that $\Pr(\mu_n(m) = w) = 1/n$. Combining this fact with equation (37), we obtain the following sequence of transformations:

$$\begin{aligned}
\Pr\left(\mu_n(m_2) = w \cap m_1 \sim_w m_2\right) &\to \frac{1}{n}\Pr(m_1 \sim_w m_2) \quad &&\text{(multiply by } n\text{)}\\
n\Pr\left(\mu_n(m_2) = w \cap m_1 \sim_w m_2\right) &\to \Pr(m_1 \sim_w m_2) &&\text{(i.i.d preferences } \forall w\text{)}\\
\sum_{i=1}^{n}\Pr\left(\mu_n(m_2) = w_i \cap m_1 \sim_{w_i} m_2\right) &\to \Pr(m_1 \sim_w m_2) &&\text{(disjoint assignments } \mu_n(w_i)\text{)}\\
\Pr\left(\bigcup_{i=1}^{n}[\mu_n(m_2) = w_i \cap m_1 \sim_{w_i} m_2]\right) &\to \Pr(m_1 \sim_w m_2) &&\text{(LHS is the desired prob)}\\
\Pr\left(m_1 \sim_{\mu_n(m_2)} m_2\right) &\to \Pr(m_1 \sim_w m_2).
\end{aligned} \tag{38}$$

This proves part (b). □

Theorem 2 *(a)The total number T_n of weak preferences over n partners satisfies the recurrence relation*

$$T_n = \sum_{i=0}^{n-1}\binom{n}{i}T_i. \tag{39}$$

(b) The total number $\tilde{T}_n$ of weak preferences over n partners in which two fixed agents are in the same indifference class satisfies the recurrence relation

$$\tilde{T}_n = T_{n-1}. \tag{40}$$

Proof. (a). The first partition block of size k $(1 \leq k \leq n)$ can be selected from the set of n partners in exactly $\binom{n}{k}$ ways, because within blocks, order does not matter. To each selection of this first k-block, there corresponds a subset of $(n-k)$ remaining elements to be partitioned into ordered blocks, which can be done in exactly T_{n-k} ways. Summing over all possible sizes $k = 1, 2, \ldots, n$ results in the sum

$$T_n = \sum_{k=1}^{n} \binom{n}{k} T_{n-k} = \sum_{k=1}^{n} \binom{n}{n-k} T_{n-k} = \sum_{i=0}^{n-1} \binom{n}{i} T_i, \tag{41}$$

where $i = n - k$.

(b). Let the set of men M be of cardinality n. If two fixed elements $m_1, m_2 \in M$ always appear in the same block of an ordered partition of M, one can treat them as a single element $\bar{m}$ for partitioning purposes. The resulting set $\{\bar{m}, m_3, m_4, \ldots, m_n\}$ consists of $n-1$ elements and can therefore be partitioned into ordered blocks in exactly T_{n-1} ways. Therefore the number of ordered partitions $\tilde{T}_n$ in which the two fixed elements m_1 and m_2 always occur in the same block, satisfies the recurrence relation $\tilde{T}_n = T_{n-1}$. □

Theorem 3 *Under the default preference-generating process $\mathscr{P}_0$, the probability that an arbitrary woman w is indifferent between two given men m_1 and m_2 satisfies the limit*

$$\lim_{n\to\infty} \Pr(m_1 \sim_w m_2\ ;\ n) = \lim_{n\to\infty} \frac{T_{n-1}}{T_n} = 0. \tag{42}$$

Proof. Since this fact can be proven in a number of ways, here I offer an exact proof not involving approximations; for a different proof based on an asymptotic approximation of the sequence T_n, see Barthélémy (1980).

Denoting $\Delta T_n \equiv (T_n - T_{n-1})$, observe that

$$\frac{T_{n-1}}{T_n} = \frac{T_{n-1}}{T_{n-1} + \Delta T_n} = \frac{1}{1 + \frac{\Delta T_n}{T_{n-1}}}. \tag{43}$$

I will prove that $\frac{\Delta T_n}{T_{n-1}} \to \infty$, which implies that $T_{n-1}/T_n \to 0$. The recurrence relation

$$T_n = \sum_{i=0}^{n-1} \binom{n}{i} T_i, \tag{44}$$

in conjunction with the fact that $T_0 = 1$ and $\binom{n}{0} = \binom{n-1}{0} = 1$, implies that

$$\Delta T_n \equiv T_n - T_{n-1} = \tag{45}$$

$$= \left[\binom{n}{1} - \binom{n-1}{1}\right] T_1 + \ldots + \left[\binom{n}{n-2} - \binom{n-1}{n-2}\right] T_{n-2} + \binom{n}{n-1} T_{n-1}. \tag{46}$$

One can simplify this using Pascal's formula $\binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}$ and the fact that $\binom{n}{n-1} = n$, obtaining

$$\Delta T_n = nT_{n-1} + \sum_{i=0}^{n-3} \binom{n-1}{i} T_{i+1}. \tag{47}$$

The last equation implies that the percentage change in T_n exceeds n in magnitude, because

$$\frac{\Delta T_n}{T_{n-1}} = n + \underbrace{\frac{\sum_{i=0}^{n-3} \binom{n-1}{i} T_{i+1}}{T_{n-1}}}_{\equiv Z(n)>0}. \tag{48}$$

Therefore,

$$\lim_{n\to\infty} \frac{\Delta T_n}{T_{n-1}} = \lim_{n\to\infty} [n + Z(n)] = \infty, \tag{49}$$

and hence,

$$\lim_{n\to\infty} \frac{T_{n-1}}{T_n} = \lim_{n\to\infty} \frac{1}{1 + \frac{\Delta T_n}{T_{n-1}}} = \frac{1}{1+\infty} = 0. \quad \square \tag{50}$$

Theorem 4 *Let α_n denote the expected number of men involved in ordered 2-tuples that admit cycles in a problem of size n, and let agent preferences be drawn according to the default preference-generating process $\mathscr{P}_0$. Then, as $n \to \infty$:*

a) *The expected number of men α_n who admit cycles increases with n:*

$$\alpha_n = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}} \;\to\; \infty. \tag{51}$$

b) *The expected proportion of men $\frac{\alpha_n}{n}$ who admit cycles converges to 0:*

$$\frac{\alpha_n}{n} = \frac{1}{2n} + \frac{1}{2n}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}} \;\to\; 0. \tag{52}$$

Proof. (a) For a given problem size n, let $Y_{i,n}$ be a random variable such that $Y_{i,n} = 1$ if the ordered 2-tuple of men indexed by i admits a cycle, and $Y_{i,n} = 0$ otherwise. Since the number of men is n, there are a total of $n(n-1)$ ordered 2-tuples of men: $i \in \{1, \ldots, n(n-1)\}$.

Now define the proportion of ordered 2-tuples admitting a cycle as

$$\pi_n \equiv \frac{1}{n(n-1)} \sum_{i=1}^{n(n-1)} Y_{i,n}. \tag{53}$$

Therefore, given a problem of size n, the expected proportion of 2-tuples admitting a cycle is

$$\begin{aligned} \mathbb{E}\pi_n =& \frac{1}{n(n-1)} \sum_{i=1}^{n(n-1)} \mathbb{E}Y_{i,n} \;=\; \frac{\cancel{n(n-1)}}{\cancel{n(n-1)}} \Pr(Y_{i,n} = 1) \;= \\ =& \Pr(m_1 \sim_{\mu(m_2)} m_2) \to \Pr(m_1 \sim_w m_2), \end{aligned} \tag{54}$$

according to Theorem 1. But as shown in Theorem 3, the probability $\Pr(m_1 \sim_w m_2)$ converges to 0. Therefore, as $n \to \infty$,

$$\mathbb{E}\pi_n \to 0. \tag{55}$$

In the same way we compute the expected *number* γ_n of ordered 2-tuples of men who admit a cycle as

$$\begin{aligned} \mathbb{E}\gamma_n = \sum_{i=1}^{n(n-1)} \mathbb{E}Y_{i,n} = n(n-1)\Pr(Y_{i,n}=1) = n(n-1)\Pr(m_1 \sim_{\mu(m_2)} m_2) \to \\ \to n(n-1)\Pr(m_1 \sim_w m_2) = n(n-1)\frac{T_{n-1}}{T_n}, \end{aligned} \tag{56}$$

according to Theorems 1 and 3.

Denote the expected number of men who participate in (ordered) 2-tuples that admit cycles with α_n. Since α_n men form exactly $\alpha_n(\alpha_n - 1)$ ordered 2-tuples, we can find the total number of men who admit cycles by solving

$$\alpha_n(\alpha_n - 1) = n(n-1)\frac{T_{n-1}}{T_n}. \tag{57}$$

The only positive root of this quadratic is

$$\alpha_n = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}}. \tag{58}$$

As it turns out, $\frac{T_{n-1}}{T_n}$ is of the order of $\frac{1}{n}$, which implies that the number of men who admit cycles actually grows with n and is of the order of $\sqrt{(n-1)}$. To demonstrate this formally, we can use Barthélémy's (1980) asymptotic approximation for the sequence T_n, from which it follows that $\frac{T_{n-1}}{T_n} \to \frac{\ln 2}{n}$. Substituting this in (58) yields

$$\alpha_n \to \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4(n-1)\ln 2} \to \infty, \tag{59}$$

which is unbounded and strictly increasing in n.

(b) Equation (58) also implies that the expected *proportion* of men who admit any cycle is

$$\frac{\alpha_n}{n} = \frac{1}{2n} + \frac{1}{2n}\sqrt{1 + 4n(n-1)\frac{T_{n-1}}{T_n}}. \tag{60}$$

Now it is easy to show that the expected proportion of men admitting a cycle, as a fraction of all men, converges to zero. The limit of α_n/n can be found as

$$\begin{aligned} \lim_{n\to\infty}\left(\frac{\alpha_n}{n}\right) = \lim_{n\to\infty}\left(\frac{1}{2n} + \sqrt{\frac{1+4(n-1)\ln 2}{4n^2}}\right) = \\ = \lim_{n\to\infty}\sqrt{\frac{1}{n^2} + \frac{n-1}{n^2}\ln 2} = \lim_{n\to\infty}\sqrt{\frac{\ln 2}{n}} = 0. \end{aligned} \tag{61}$$

Theorem 5 *If a stochastic preference-generating process for weak preferences $\mathscr{P}^*$ is symmetric and agent preferences are mutually independent, then the following statements hold:*

a) *If the probability of a pairwise indifference* $\Pr(m_1 \sim_w m_2)$ *under $\mathscr{P}^*$ is of the order of $1/n^2$, then $\alpha_n \to const.$ and $\frac{\alpha_n}{n} \to 0$ as $n \to \infty$.*

b) *If the probability of a pairwise indifference* $\Pr(m_1 \sim_w m_2)$ *under $\mathscr{P}^*$ is of the order of $1/n^k$ for $k \geq 3$, then $\alpha_n \to 0$ and $\frac{\alpha_n}{n} \to 0$ as $n \to \infty$.*

Proof. To prove this result, replace the expression T_{n-1}/T_n with the expression n^{-k} in formulas (60) and (58) and take the limit $n \to \infty$ to obtain

$$\begin{aligned} \frac{\alpha_n}{n} &= \frac{1}{2n} + \sqrt{\frac{4n^{1-k}(n-1)}{4n^2}} \to \sqrt{n^{-k}} = n^{-k/2} \to 0 \quad (\forall k > 0); \\ \alpha_n &= \frac{1}{2} + \frac{1}{2}\sqrt{1 + 4n^{1-k}(n-1)} \to \begin{cases} \to (1+\sqrt{5})/2 & \text{if } k = 2 \\ \to 1 & \text{if } k \geq 3 \end{cases}. \end{aligned} \tag{62}$$

The result $\alpha_n \to 1$ obtained for $k \geq 3$ implies that, on average, only one man will admit a cycle when $n \to \infty$. However, recalling from the definition of a cycle that a man can never form a cycle alone (without envying another agent), in reality this implies that there are no 2-tuples that form cycles; thus we must round the expected number of participants in such 2-tuples down to zero: $\alpha_n \to 0$.

Notice the principal difference with the case where $k = 2$ and $\alpha_n \to (1+\sqrt{5})/2$. Although the limit $(1+\sqrt{5})/2 \approx 1.62$ is also a small number, it rounds up to 2, the smallest possible number of cycle participants. Therefore the difference between the case where $k = 2$ and $k = 3$ is that in the former case, one (minimal) cycle is admitted in expectation, while in the latter case, there are none. □

6.2 Extension 1: Pareto-improvement Chains

Here I drop the assumption that all agents are acceptable. Partners who appear as choices worse than i are *unacceptable* for i, and if $\mu(i) = i$ we say that agent i remains single under the matching μ. If an agent remains single, he or she cannot be part of a Pareto-improvement cycle, but may still be part of a Pareto-improvement chain.

In a setting with single agents, a man can point to either a matched man or a single woman, so instead of 2-tuples of men of the type $m_i \rightarrow m_j$, now I will consider generalized envy 2-tuples of the kind $m \rightarrow i$ where the envied agent i can be either a matched man or a single woman.

Given a matching μ, denote the set of matched men as $\overline{M}$ and the set of single women as $\underline{W}$; I consider 2-tuples of the type (m, i) so that $i \in \overline{M} \cup \underline{W}$. Note that for every single man there is exactly one single woman, so $|\overline{M}| + |\underline{W}| = n$.

Lemma 4 *(Necessary condition for Pareto-improvement cycles and chains.) Suppose that μ is a stable matching. Then the 2-tuple of agents $m \in M$ and $i \in \overline{M} \cup \underline{W}$ does not admit a Pareto-improvement cycle or Pareto-improvement chain unless*

$$m \sim_{\mu(i)} i \quad \text{or} \quad m \sim_{\mu(m)} i, \ \text{or both.} \tag{63}$$

Proof. First suppose that $\mu(i) P_m \mu(m)$, where $m \in M$ and $i \in \overline{M} \cup \underline{W}$. (If instead $\mu(m) P_m \mu(i)$, agents m and i will never voluntarily trade partners.)
1. *Proof for cycles*: Suppose not, so $m \not\sim_{\mu(i)} i$. Then either $\mu(m_2)$ will be unwilling to trade, or else $(m_1, \mu(m_2))$ form a blocking pair, so μ is not stable.
2. *Proof for chains:* If both m and i are matched men, the proof is the same as for cycles. If $m \in \overline{M}$ (so m has a partner) and $i \in \underline{W}$ (i.e., i is a single woman), then obviously $\mu(i) = i$. If i is not indifferent between m and remaining single, either she will not agree to trade, or else (m, i) are a blocking pair. Alternatively, if m is a single man, then i is a matched man, and if $m \not\sim_{\mu(i)} i$, again $\mu(i)$ will either refuse to trade, or form a blocking pair with m.
It is also possible that $\mu(m) P_i \mu(i)$; in this case, simply reverse the roles of m and i in the proofs above to obtain the second part of the necessary condition $m \sim_{\mu(m)} i$. (If instead $\mu(i) P_i \mu(m)$, agents m and i will not want to trade.) □

The main difference from the baseline case is that now each woman's preference is defined over $n+1$ agents: the n men plus herself. As before, we will say as a matter of convention that the ordered 2-tuple (m, i) admits a cycle if $m \sim_{\mu(i)} i$. Therefore we have

$$\Pr(m \sim_{\mu(i)} i) \rightarrow \Pr(m \sim_j i) = \frac{T_n}{T_{n+1}} \rightarrow 0 \tag{64}$$

because the problem is now of size $n+1$. Since $|\overline{M}| + |\underline{W}| = n$, the total possible number of ordered envy 2-tuples is $n(n-1)$ as before, so the remainder of the proofs in Theorem 4 carry over without change.

6.3 Extension 2: Weak Preferences for Men and Women

Similar results hold when both men and women have weak preferences, except that in this case the inefficiency vanishes slower. Given a market size n, now each man and woman independently draws *weak* preferences from the uniform distribution over all possible weak preferences over the n agents of the opposite side. The draw is repeated for each market size n; successive draws are independent. In this setting, cycles can occur more frequently, because there are more indifferent agents who can agree to switch partners. The necessary condition for a cycle involving two fixed men m_1 and m_2 is now the following.

Lemma 5 *Suppose that μ is a stable matching. Then two fixed men m_1 and m_2 do not admit a Pareto-improvement cycle unless at least one of the following four conditions holds:*

$$m_1 \sim_{\mu(m_1)} m_2; \quad m_1 \sim_{\mu(m_2)} m_2; \quad \mu(m_1) \sim_{m_1} \mu(m_2); \quad \mu(m_1) \sim_{m_2} \mu(m_2). \tag{65}$$

Proof. The first two indifference conditions, $m_1 \sim_{\mu(m_1)} m_2$ and $m_1 \sim_{\mu(m_2)} m_2$, follow without change from the proof of Lemma 2. In addition, however, now a cycle can also occur when one of the women $\mu(m_1), \mu(m_2)$ has a *strict* preference over the men m_1, m_2, but one of these men is indifferent between $\mu(m_1)$ and $\mu(m_2)$. Again, if $m_1 P_{\mu(m_1)} m_2$, then $\mu(m_1)$ and $\mu(m_2)$ will not trade, and when $m_2 P_{\mu(m_1)} m_1$, the two fixed men cannot be in the same cycle unless $\mu(m_1) \sim_{m_2} \mu(m_2)$; the same logic applies to $\mu(m_1)$, resulting in the indifference $[\mu(m_1) \sim_{m_1} \mu(m_2)]$. □

Theorem 6 *Let μ be a random stable matching selected by the RSM and let preferences for both men and women be randomly drawn according to a preference-generating process $\mathscr{P}_0'$, whereby both men and women have weak, uniformly drawn preferences as in Definition 2(b). Then the probability that two arbitrary fixed men m_1 and m_2 admit a Pareto-improvement cycle satisfies the limit*

$$\Pr\big((m_1, m_2) \text{ admit cycle}\big) \to 0 \text{ as } n \to \infty. \tag{66}$$

Proof. When both sides of the market have weak preferences, I will say that the ordered 2-tuple (m_j, m_k) admits a cycle if and only if the event $[m_j \sim_{\mu(m_k)} m_k] \cup [\mu(m_j) \sim_{m_k} \mu(m_k)]$ occurs.

The probability that an arbitrary ordered pair of men (m_1, m_2) admits a cycle is

$$\begin{aligned}\Pr\big((m_1, m_2) \text{ admit cycle}\big) &= \Pr\big([m_1 \sim_{\mu(m_2)} m_2] \cup [\mu(m_1) \sim_{m_2} \mu(m_2)]\big) \leq \\ &\leq \Pr\big(m_1 \sim_{\mu(m_2)} m_2\big) + \Pr\big(\mu(m_1) \sim_{m_2} \mu(m_2)\big) = \\ &= 2\Pr(m_1 \sim_{\mu(m_2)} m_2).\end{aligned} \tag{67}$$

But according to Theorems 1(b) and 3, we have

$$2\Pr(m_1 \sim_{\mu(m_2)} m_2) \to 2\Pr(m_1 \sim_w m_2) = 2\frac{T_{n-1}}{T_n} \to 0, \tag{68}$$

which implies that $\Pr\big((m_1, m_2) \text{ admit cycle}\big) \to 0$. $\square$

Having established this fact, the remaining proofs (Theorem 4) apply directly without change, except that the rate of convergence to zero is twice larger. This completes the extension to weak preferences for both sides of the market.

References

[1] Abdulkadiroğlu, A., Pathak, P., Roth, A. E.: Strategy-Proofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match. American Economic Review **99**(5), 1954-1978 (2009).

[2] Ashlagi, I., Kanoria, Y., Leshno, J.D..: Unbalanced Random Matching Markets: The Stark Effect of Competition. Working paper (2015).

[3] Barthélémy, J.-P.: An Asymptotic Equivalent for the Number of Total Preorders on a Finite Set. Discrete Mathematics **29** (3), 311-313 (1980).

[4] Che, Y., Kojima, F.: Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms. Econometrica **78**(5), 1625–1672 (2010).

[5] Erdil, A., Ergin, H.: What's the Matter with Tie-Breaking? Improving Efficiency in School Choice. American Economic Review, **98**(3), 669-689 (2008).

[6] Erdil, A., Ergin, H.: Two-Sided Matching with Indifferences. Working Paper (2015).

[7] Gusfield, D., and Irving, R.W.: The Stable Marriage Problem: Structure and Algorithms. Cambridge and London: MIT Press (1989).

[8] Immorlica, N., Mahdian, M.: Marriage, Honesty, and Stability. Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 53-62 (2005).

[9] Kojima, F., Manea, M.: Incentives in the Probabilistic Serial Mechanism. Journal of Economic Theory 145, 106–123 (2010).

[10] Kojima, F., Pathak, P.: Incentives and Stability in Large Two-Sided Matching Markets. American Economic Review **99**(3), 608-627 (2009).

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

[12] Roth, A. E.: A Natural Experiment in the Organization of Entry-Level Labor Markets: Regional Markets for New Physicians and Surgeons in the United Kingdom. American Economic Review **81**(3), 415–440 (1991).

[13] Roth, A. E., Peranson, E.: The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design. American Economic Review **89**(4), 748–780 (1999).

[14] Roth, A. E., Vande Vate, J. H.: Random Paths to Stability in Two-Sided Matching. Econometrica **58**(6), 1475–1480 (1990).

[15] Roth, A. E., Sotomayor, M.: Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis. Cambridge and New York: Cambridge University Press (1990).