LARGE ROOMMATE PROBLEM WITH NON-TRANSFERABLE RANDOM UTILITY

MARCIN PĘSKI

ABSTRACT. We analyze a large roommate problem (i.e., marriage matching in
which the marriage is not restricted solely to matchings between men and women)
with non-transferable utility. It is well known that while a roommate problem may
not have a stable proper matching, each roommate problem does have an stable
improper matching. In a random utility model with types from Dagsvik (2000) and
Menzel (2015), we show that all improper stable matchings are asymptotically close
to being a proper stable matching. Moreover, the distribution of types in stable
matchings (proper or not) converges to the unique maximizer of an expression that
is a sum of two terms: the average "welfare" of the matching and the Shannon
entropy of the distribution. In the noiseless limit, when the random component of
the utility is reduced to zero, the distribution of types of matched pairs converges
to the outcome of the transferable utility model.

1. INTRODUCTION

We analyze stable matchings in the large "roommate" model (Gale and Shapley
(1962)) with random and non-transferable utility. We assume that the utility of
individual i matching with individual j is drawn from a distribution

$$U_i^j \sim F_{t_i}^{t_j, \theta_{\{i,j\}}} \tag{1}$$

that depends on the types t of both of individuals (for example, income, beauty, etc.),
and an unobservable match-specific shock $\theta_{\{i,j\}}$. The match-specific shock is drawn
from a distribution that depends on the types of the individuals. The key assumption
is that distributions $F_t^{s,\theta}$ are absolutely continuous with respect to each other at the

top end of their supports. The assumption ensures that the utilities are subject to a non-trivial randomness that remains even after learning the types and match-specific shock of two individuals. A special case is the random utility specification of Dagsvik (2000):

$$U_i^j = \upsilon_{t_i}^{t_j} + \gamma \varepsilon_i^j, \tag{2}$$

where $\upsilon_{t_i}^{t_j}$ is the deterministic part of the utility, the random shock ε_i^j is independently distributed with extreme value type I distribution, and $\gamma > 0$ is a parameter.

The roommate problem is a generalization of the classic marriage matching model in which matching is no longer restricted solely to members of two different sides of the market. In the marriage literature, the roommate problem allows to study gay marriage, including marriage markets in which the participants choose between two types of marriage. In the job-worker matching, the roommate problem allows the workers to form partnerships. An important application of the roommate problem is kidney exchange, where an "individual" is composed from two people, donor and patient. Using only the compatibility of the blood type, Roth, Sonmez, and Unver (2007) argue that 3-way exchange cycles are enough to maximize the number of compatible matchings. (Here, 3-way cycle corresponds to a cycle composed of 3 badly matched individuals in our terminology). Their result does not necessarily apply when the quality of the match is affected also by the tissue as well as blood compatibility. If we interpret the ε errors as the tissue compatibility, our asymptotic result indicates that two-way exchanges are sufficient to implement stable matching for a great majority of the individuals. [1]

It is well-known that the roommate problem may not have a stable proper matching. For this reason, we use an approximate concept of (improper) matching where some are (badly) matched with two, instead of one, other individuals. The concept was introduced in Tan (1991) who established the existence of a stable (but, possibly, improper) matching.

Our main result establishes two asymptotic properties of *all* stable (and, possibly, improper) matchings when the population grows large and utilities are derived from the random utility model (1). First, we show that, with a probability converging to

---

[1] I am very grateful to the anonymous referee for this suggestion.

1, the fraction of badly matched individuals in any stable matching converges to 0. It follows that all stable matchings are *asymptotically proper.* Second, we characterize the limit of the probability masses $d \in \Delta \left(T \cup \{\emptyset\} \times T \cup \{\emptyset\} \right)$ over the pairs of types of matched individuals. Here, match with $\emptyset$ corresponds to staying alone. As the size of the population grows to infinity, the distributions over the matched types in any stable matching converge to the unique solution of the concave maximization problem

$$\max_{d} \; \text{Welfare} \left(d; F^{\cdot}_{\cdot} \right) + \frac{1}{\gamma} \text{Entropy} \left(d \right) \; \text{ st. } d \text{ satisfies feasibility restrictions.} \qquad (3)$$

The welfare term of formula (3) is equal to the average sum of "deterministic utilities" in the stable match, where the "utility" coefficients are derived from the top end densities of distributions (1). The second term does not depend on the properties of the distributions (1) and is a version of a well-known measure of disorder from thermodynamics (and it also bears resemblance to the Shannon measure of informational content). As in physics, the entropy here has a counting interpretation and it is equal (modulo a constant) to the logarithm of the number of micro-configurations (i.e., individual matchings) that are consistent with a given macro-state (i.e, distribution d). In particular, the larger the entropy, the more imprecise the information about the individual matchings. Finally, the most important feasibility restriction is that the marginals of distribution d match the distribution of types in the population. Other feasibility restrictions may be required to reflect special additional features of the environment. For example, in case of the marriage matching, "male" types are restricted from marrying other "male" types, etc. Together, our result says that the stable matching balances the "welfare" maximization with an entropy cost.

The random utility model (1) was introduced in Dagsvik (2000) for the case of the marriage matching (and extreme value type 1 distribution of shocks). Dagsvik (2000) relies on a heuristic to argue that the limit of stable distributions must be the unique solution to a system of demand-supply type of equations. Menzel (2015) provides a formal proof of the claim in Dagsvik (2000). The ingenious argument follows the demand-supply logic and it relies on various probability bounds obtained from a careful analysis of the rejection chains in the Gale-Shapley algorithm.

This paper contributes to the literature in the following ways. First, we extend the earlier results to the roommate problem. Because the Gale-Shapley algorithm does not work in the roommate problem, the Menzel (2015) proof does not apply and our proof is necessarily very different. Our argument is direct and it consists of two parts. In the first part, we estimate a combinatorial upper bound on the number of all (not necessarily stable) matchings that induce a given distribution d. It turns out that this number is increasing with entropy. In the second part, we estimate an upper bound on the probability that a given matching is stable. We show that this probability is monotonically increasing with the "average welfare" term of (3). The estimate relies on approximations from the large deviations theory (see Ellis (2005)). By multiplying both estimates, we obtain an upper bound on the probability of a stable matching with type distribution d. We show that if d does not maximize expression (3), then the probability of a stable matching with a type d converges to 0 at an exponential rate. Thus, the distribution that is induced by all stable matching must be a maximizer of (3).

Second, Menzel (2015) assumes that the utility shocks are independent across partners in the match. However, utility correlations are common in a typical matching situation.[2]) Examples of common shocks include:

- random events that affect both parties simultaneously (weather during a first date, market conditions during a job interview),
- common history (attending the same university, common friends),
- unobserved relationship characteristics (part-time vs full-time, possibility of working from home, etc.) that jointly affect the preferences of both matching parties.

It is important to understand how the correlated shocks affect the predictions of the model. For this reason, we allow for any common unobservable match-specific shocks. Our proof is simple and flexible enough so that the match correlations do not cause any extra complication.

Third, Dagsvik (2000) and Menzel (2015) show that the unique limit distribution is a solution to a system of demand-supply type of equations. Our result is consistent

---

[2]Dagsvik (2000) allows for correlations in his model, but his argument remains heuristic.

with their findings, as the unique maximizer of (3) solves the same set of equations. We are not aware of any other, earlier use of formula (3) to describe the social outcomes. Both approaches are complementary and shed light on different aspects of the problem. In particular, our approach emphasizes the fact that the likelihood of a certain social outcome is directly related to its "randomness", or the cardinality of micro-configurations that are consistent with the observed outcome.

Finding the exact solution to the maximization problem (3) is not easy, as it requires solving a system of non-linear first-order conditions. However, our characterization implies that any computational method that behaves well in the concave maximization problem can be readily used to find the approximate solutions.

Fourth, formula (3) leads to a surprising connection between large population matching with *non-transferable* and random utility (2) and matching models with *transferable* utility with coefficients vs. We consider a noiseless limit of the model (2) where $\gamma \to 0$. In the noiseless limit, expression (3) becomes dominated by the first term and the stable matching distribution maximizes average welfare. The latter is also the unique distribution obtained in matching with preferences given by vs and *transferable* utility. At the same time, the limit is typically *not* equal to the distribution obtained when utility is non-transferable and $\gamma = 0$. We interpret the result as a non-transferable foundation for the transferable utility model.

Choo and Siow (2006) introduced a different version of marriage matching model (2) with transferable utility and where the random shock $\varepsilon_i^j = \varepsilon_i^{t_j}$ is the same for all individuals j of the same type. Working with that model, Galichon and Salanie (2012) show that, with the transferable utility, the limit distribution over types in stable matching with transfers (i.e., welfare maximizing matching) converges to the solution of a version of (3), but where Entropy (d) is replaced by a "generalized entropy". The "generalized entropy" is defined relative to the distribution of the vector of the shocks $(\varepsilon_{\cdot}^t)_{t \in T}$ and it is equal to the value of a certain stochastic allocation problem. The main difference with our paper is that the entropy in (3) does not depend on the distribution of the shocks and the entire impact of the distribution is captured by the "average welfare" term.

Section 2 describes a general roommate problem. Sections 3 and 4 introduce, respectively, the models of types and random utility. Section 5 presents the main result and describes the main ideas of the proof. Section 6 discusses the noiseless limit and the connection to the transferable utility model. Section 7 uses heuristic argument in a special case of the model to illustrate the main ideas behind the proof.

2. Roommate problem

In this section, we discuss the roommate problem. There are N individuals $i \in I$. An individual i may form a match with another individual $j \neq i$, in which case i receives utility U_i^j. Alternatively, individual i may decide to remain alone with utility U_i^i.

We allow for the possibility that some matches are not feasible. Let $R \subseteq I \times I$ be the set of restricted matches. We assume that the restriction is symmetric: if $(i, j) \in R$, then $(j, i) \in R$, and that the option of staying alone is not restricted: $(i, i) \notin R$. For example, in the marriage matching model of Gale and Shapley (1962), the population is divided into two sexes, and only opposite-sex matches are allowed.[3]

A *proper matching* is a bijection $\mu : I \to I$ such that $\mu(\mu(i)) = i$ for each i and such that $(i, \mu(i)) \notin R$. Proper matching μ is *stable* if

$$\forall_{(i,j) \notin R} \text{ either } U_i^j \leq U_i^{\mu(i)} \text{ or } U_j^i \leq U_j^{\mu(j)} \tag{4}$$

In other words, a matching is stable if there is no unrestricted blocking match.

It is well-known that a roommate model may not have a stable proper matching (and, in the particular, that the Gale-Shapley algorithm does not apply)[4]. Instead, we use a relaxed solution proposed in Tan (1991). An *(improper) matching* is a

---

[3]An alternative way of modeling match restrictions is to assume that the utility from restricted matches are $-\infty$, or at least smaller than the utility from staying alone. We take the explicit approach because it is more convenient when we present the random utility model.

[4]The issue can be illustrated with a simple example from Gale and Shapley (1962). There are three individuals $i = 1, 2, 3$, with preferences $U_1^2 > U_1^3 > U_1^1, U_2^3 > U_2^1 > U_2^2, U_3^1 > U_3^2 > U_3^3$. Any proper matching can be blocked by a couple that includes the remaining individual. Using the definition below (and inequality (5) in particular), one can check that there is a unique stable and improper matching: $\mu(1) = 3, \mu(2) = 1$, and $\mu(3) = 2$.
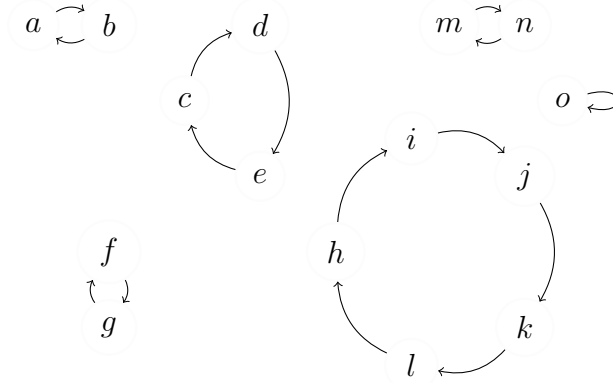
FIGURE 1. An (improper) matching.

bijection $\mu : I \to I$ such that $(i, \mu(i)) \notin R$.[5] If $\mu(\mu(i)) = i$, then, we say that i is properly matched; otherwise, i is badly matched. Because μ is a bijection, all badly matched individuals can be divided into cycles $i, \mu(i), \mu^2(i), ..., \mu^n(i) = i$, where each individual is matched with a predecessor and successor in the cycle and all members of the cycle are badly matched. An example of a matching is drawn on Figure 1. An arrow from individual i to j indicates that $\mu(i) = j$. Thus, individuals a, b, f, g, m, n are properly matched, individuals c, d, e, h, i, j, k, l are badly matched, and individual o remains single.

For each matching μ, define a mapping $c^\mu : I \to \{s, p, b\}$ that assigns each individual i with a category of its match:

$$
c^\mu(i) = \begin{cases} s, & \text{if } \mu(i) = i \text{ (i.e., if } i \text{ is single),} \\ p, & \text{if } \mu(i) \neq i, \mu(\mu(i)) = i, \text{ (ie., if } i \text{ is properly matched),} \\ b, & \text{otherwise.} \end{cases}
$$

(Improper) matching μ is *stable* if (4) holds and, additionally, for each i,

$$
U_i^{\mu(i)} \leq U_i^{\mu^{-1}(i)}. \tag{5}
$$

---

[5]Tan (1991) refers to μ as a partition.

The first requirement is obvious. The second requirement bites only for the badly matched individuals and it plays the role of sorting: each badly individual is matched with two other individuals, and $\mu(i)$ is the worse of two matches. The role of the second requirement is to eliminate equivalent matchings (for example, with the matchings with the reversed arrows in Figure (1)) and it simplifies the combinatorial analysis below. One can eliminate the second requirement and all the results go through with minor but tedious modifications to the proofs.

Assumption 1. *(Strict Preferences) For each i and $j \neq j'$, $U_i^j \neq U_i^{j'}$.*

The assumption is satisfied with probability 1 in the random utility model described later in the paper.

Theorem 1. *(Tan (1991)) Assume Strict Preferences. Then, there exists a stable (possibly, improper) matching μ.*[6]

3. TYPES

Following Dagsvik (2000) and Choo and Siow (2006), we are interested in the distribution of types among the matched individuals. A type is any observable characteristic, including gender, beauty, income, education. A type of individual i is denoted as $t_i \in T$, where T is a finite set.[7] Let $\alpha_t = \frac{1}{N}\{i : t_i = t\}$ denote the fraction of individuals with type t.

We assume that the set of restrictions R is *type-measurable*: ie., there is $R^* \subseteq T \times T$ such that for each $i \neq j$, $(i,j) \in R$ if and only if $(t_i, t_j) \in R^*$. There are two special cases:

- *no-restriction case*, in which $R^* = \emptyset$, and

---

[6]Theorem 1 is closely related to Chiappori, Galichon, and Salanie (2014) who consider the roommate problem with transferable utility. They show that if each individual can be cloned into two identical copies (with the same preferences, and treated in the same way by other players), then there exists a proper stable matching.

[7]This assumption can be relaxed to a situation where T is compact. The idea is to divide ("discretize") the space of types T into finitely many small areas. The formal argument would not add much to the difficulty of the current proof, but it would complicate the exposition. We will provide the details for an interested reader.

- the Gale-Shapley *marriage matching*, where the set of types $T = M \cup W$ is a disjoint union of men types M and women types W and $R^* = M \times M \cup W \times W$.

A *type distribution* is a probability measure $d \in \Delta\left(T \times T \times \{s, p, b\}\right)$. Each matching μ induces a matching type d^μ so that for each (t, t', c),

$$d^\mu\left(t, t', c\right) = \frac{1}{N}\left\{i : t_i = t, t_{\mu(i)} = t', c^\mu\left(i\right) = c\right\}.$$

In other words, d^μ is a probability distribution of an ordered tuple of random variables (T, T', C), where T is a type of a randomly drawn individual i in the population, T' is the type of $\mu\left(i\right)$, and C is the category of the match $(i, \mu\left(i\right))$. We write $d\left(c\right) = \sum_{t,t'} d\left(t, t', c\right)$ for each $c \in C$. Let $D^N\left(\alpha\right)$ denote the set of type distributions d^μ induced from some matching μ.

Any induced type distribution $d = d^\mu$ must satisfy the following equalities for each t:

$$\sum_{t',c} d\left(t, t', c\right) = \alpha_t, \tag{6a}$$

$$d\left(t, t', c\right) = 0 \text{ for each } t' \text{ st. } (t, t') \in R^* \text{and each } c, \tag{6b}$$

$$d\left(t, t', s\right) = 0 \text{ for each } t' \neq t, \tag{6c}$$

$$d\left(t, t', p\right) = d\left(t', t, p\right) \text{ for each } t', \tag{6d}$$

$$\sum_{t'} d\left(t, t', b\right) = \sum_{t'} d\left(t', t, b\right) =: \beta_t\left(d\right), \tag{6e}$$

Identity (6a) ensures that all type t individuals are counted. Identity (6b) reflects the restrictions. Identity (6c) is about single matched individuals. Identity (6d) is implied by the fact that all proper matches are symmetric. Identity (6e) follows from the fact that each bad match of individual i with $\mu\left(i\right)$ corresponds to exactly one bad match with individual $\mu^{-1}\left(i\right)$. We define $\beta_t\left(d\right)$ as the fraction of badly matched individuals of type t.

For each $\alpha \in \Delta T$ and $(\beta_t) \in [0, 1]^T$, let $D\left(\alpha, (\beta_t)\right)$ be the set of matching types that satisfy (6a)-(6e) and such that $\beta_t\left(d\right) = \beta_t$ for each t. Finally, let $D\left(\alpha\right) = D\left(\alpha, \mathbf{0}\right)$ be the subset of matching types of proper matchings (ie., ds such that $d\left(b\right) = 0$). Sometimes, we shall use the following assumption:

Assumption 2. *Set* $\{d \in D\left(\alpha\right) : d\left(s\right) = 0\}$ *is non-empty.*

If the assumption is satisfied, then one can always divide the individuals into pairs (with the possible exception of insignificant fractions of the population). The assumption is always satisfied in the no-restriction case. In the marriage matching case, the assumption is satisfied if and only if the number of men and women are equal, $\sum_{t \in M} \alpha_t = \sum_{t \in W} \alpha_t$.

For each matching type d, define

$$\text{Entropy}(d) = -\sum_t d(t,t,s) \log d(t,t,s) - \frac{1}{2} \sum_{t,t'} d(t,t',p)(\log d(t,t',p) + 1) \quad (7)$$
$$- \sum_{t,t'} d(t,t',b)(\log d(t,t',b) + 1).$$

We refer to Entropy (d) as a "per capita entropy" of d. The coefficient $\frac{1}{2}$ in front of the proper pairwise terms comes from the fact that each such match involves two people. Because each badly matched individual belongs to two matches, there is no analogous coefficient in front of the "entropy" of bad matches.

The per-capita entropy is a normalized version of the Shannon entropy of random variable (T, T'), where T is a type of randomly drawn individual and T' is the type of his or her match. The Shannon entropy is typically interpreted as a measure of "randomness" (or "unpredictability") of distribution d. In this paper, the role of the entropy comes from a closely related fact: it serves as a bound on the number of matchings of a given type. For each matching market of size N, let

$$\mathcal{M}_N(d) = \{\mu : d^\mu = d\}.$$

Lemma 1. *For each matching type $d \in D(\alpha, (\beta_t))$*

$$\frac{1}{N} \log |\mathcal{M}_N(d)| \quad (8)$$
$$\leq \left(\frac{1}{2} d(p) + d(b)\right) \log N + \sum_t \alpha_t \log \alpha_t + \sum_t \beta_t \log \beta_t + \text{Entropy}(d) + o(1)$$

The reminder term $o(1)$[8] does not depend on d.

---

[8]Recall that $o(1)$ denotes a sequence $a_N \to 0$ as $N \to \infty$.

The proof can be found in Appendix A. [9] Some of the ideas behind the Lemma are described using an example in Section 7

4. Random utility

For each $i \neq j$, the utility U_i^j is conditionally independently drawn from distribution (1), that depends on the type of both of the individuals, and an unobservable match-specific shock $\theta_{\{i,j\}} \in \Theta$. For each i, utility U_i^i is drawn from distribution F_{t_i}.

Following Menzel (2015), we assume that utility U_i^i is equal to the maximum among J_N i.i.d draws from distribution F_{t_i} and that $\lim N^{-1/2} J_N = j^* \geq 0$. We allow that $j^* = 0$. If $j^* > 0$, this assumption ensures that, as $N \to \infty$, the fraction of the population that remains single does not converge to 0. If $j^* = 0$ (for instance, because $J_N = 1$ for each N), each individual receives many more utility draws from pairwise matches than from staying alone, which leads to asymptotically small fractions of populations who prefer to stay single. If J_N grows at rate higher than $N^{-1/2}$, then almost all individuals remain single in large populations.

The match-specific shock is chosen independently for each match from probability distribution $\Phi^{t,t'}$ over finite set Θ. Its role is to generate (unobservable) correlations between the utilities of match partners. The existence of such correlations is natural in almost all matching situations.

Definition 1. The random utility model $\mathcal{F} = \left\{ F_t^{t',\theta}, F_t \right\}$ is *regular* if each distribution $F \in \mathcal{F}$ (a) is atomless, (b) it has support restricted to $[0, 1]$, (c) it has continuous density f with respect to the Lebesgue measure on the interval, and (d) the top end density is strictly positive $f(1) > 0$.

Part (d) is the key restriction: it ensures that, at least at the top of the support of the distributions, even after learning the type and the match specific shocks, the utility is subject to non-trivial remaining randomness. Our characterization of matching type in stable matchings depends on the distribution of utilities only through the top end derivatives.

---

[9]Although it is not necessary for our purposes, one can show that the bound (8) is asymptotically tight.

Because the notion of stability is purely ordinal, the regularity assumption can be relaxed in the following way. The model is *essentially regular* if there exist strictly increasing functions $\phi_t : R \to R$ for each type t such that the random utility model $\mathcal{F}' = \left\{ F_t^{t',\theta} \circ \phi_t, F_t \circ \phi_t \right\}$ is regular. Example 1 below illustrates this definition. From now on, we assume that the model is regular.

For any two types t, t', θ, let

$$V_t^{t',\theta} := f_t^{t',\theta} (1) > 0, V^{t,t'} = \sum_\theta \Phi^{t,t'} (\theta) \, V_t^{t',\theta} V_{t'}^{t,\theta}, \text{ and } \upsilon^{t,t'} = \log V^{t,t'}, \qquad (9)$$

$$V_t := f_t (1) > 0, \text{ and } \upsilon^t = \log V_t + \log j^*.$$

We refer to $\upsilon^{t,t'}$ as the "welfare" of a match of individual type t with type t' and υ^t as the "welfare" of staying single. The latter is modified by the normalized number of draws j^* that is used to determine the utility of being single.

We use the term "welfare" for two reasons. First, $\upsilon^{t,t'}$ has a natural interpretation of deterministic (or, observable) welfare in some applications (see Example 1 below). Second, constant $\upsilon^{t,t'}$ plays the role of match welfare in the following formula. For each matching d, let

$$\text{Welfare}\,(d; \upsilon.) = \sum_t d\,(t, t, s)\,\upsilon^t + \frac{1}{2} \sum_{t,t'} d\,(t, t', p)\,\upsilon^{t,t'} + \sum_{t,t'} d\,(t, t', b)\,\upsilon^{t,t'}. \qquad (10)$$

We interpret $\text{Welfare}\,(d; \upsilon.)$ as a average "welfare" of matching type d. The coefficient $\frac{1}{2}$ in front of the average welfare of proper pairwise matches comes from the fact that each such match involves two people. Because each badly matched individual belongs to two matches, there is no analogous coefficient in front of average "welfare" in bad matches. By convention, the value of (10) is well-defined whenever $j^* > 0$ or $d\,(s) = 0$; otherwise, it is equal to $-\infty$.

Example 1. Suppose that $\upsilon_t^{t'}$ are constants and suppose that the random utility is the sum

$$u_i^j = \upsilon_{t_i}^{t_j} + \varepsilon_i^j$$

of a type-dependent deterministic (normalized) utility $\upsilon_i^{t_j}$, and an idiosyncratic random shock $\varepsilon_i^j \geq 0$ that is drawn i.i.d from distribution G. Further, suppose that the

support of the distribution G is unbounded, and for each a, the limit

$$\lim_{x \to \infty} \frac{1 - G(x + a)}{1 - G(x)} > 0 \tag{11}$$

exists and is strictly positive. (Examples include the extreme value type I distribution used in Dagsvik (2000) and the exponential distribution.) This model is essentially regular with welfare constants equal to $v^{t,s} = v_s^t + v_t^s$. Indeed, let $U_i^j = G\left(u_i^j\right)$. Then, the cumulative distribution function F_i^j of U_i^j is equal to $F_i^j(x) = G\left(G^{-1}(x) - v_{t_i}^{t_j}\right)$.[10] Moreover, one can show that the logarithm of the density at 1 is equal to[11]

$$\log f_i^j(1) = C_G v_{t^i}^{t_j}$$

for some constant $C_G > 0$.

5. MAIN RESULT

The main result characterizes the matching type of stable matchings. Consider a maximization problem

$$\max_{d \in D(\alpha)} \text{Welfare}(d; v) + \text{Entropy}(d). \tag{12}$$

Problem (12) has a solution if and only if the objective function is well-defined over at least some part of the feasible set, or when either $j^* > 0$, or Assumption 2 is satisfied. Moreover, because Welfare $(d; v)$ is linear and Entropy (d) is strictly concave, if the solution exists, it is necessarily unique. We denote the solution as $d^*(\alpha, v)$.

The constraint $d \in D(\alpha)$ ensures that the solution respects the distribution of types, feasibility restrictions R^*, and, importantly, assigns probability 1 to proper

---

[10]$F_i^j(x) = P\left(U_i^j < x\right) = P\left(G\left(u_i^j\right) < x\right) = P\left(u_i^j < G^{-1}(x)\right) = P\left(\varepsilon_i^j < G^{-1}(x) - v_{t_i}^{t_j}\right) = G\left(G^{-1}(x) - v_{t_i}^{t_j}\right).$

[11]Let $L(v)$ denote the value of the limit in (11). It must be that $L(v + v') = L(v) L(v')$, or $L(v) = e^{-C_G v}$ for some constant C_G. Because the limit is decreasing, it must be that $C_G > 0$. By L'Hospital's rule,

$$f_i^j(1) = \lim_{y \to 1} \frac{d}{dy}\left(G\left(G^{-1}(y) - v_{t_i}^{t_j}\right)\right) = \lim_{x \to \infty} \frac{g\left(x - v_{t_i}^{t_j}\right)}{g(x)} = \lim_{x \to \infty} \frac{1 - G\left(x - v_{t_i}^{t_j}\right)}{1 - G(x)} = e^{C_G v_{t_i}^{t_j}}.$$

matches. In particular, because $d(b) = 0$ for each $d \in D(\alpha)$, problems (3) and (12) are equivalent.

Theorem 2. *Fix a type space T, a measure $\alpha \in \Delta T$, a match shock space Θ, distributions $\Phi^{t,t'}$ of match shocks, and a regular random utility model $\left\{F_t^{t',\theta}, F_t\right\}$. Let $\upsilon^{\cdot}$ be derived as in (9). Suppose that either $j^* > 0$, or Assumption 2 holds. Then, there exists a unique solution $d^*(\alpha, \upsilon)$ to problem (12) and*

$$\lim_{N \to \infty} Prob_N\left(\|d^\mu - d^*(\alpha, \upsilon)\| < \varepsilon \text{ for each stable matching } \mu\right) = 1.$$

The proof of Theorem 2 can be found below. The theorem says that for sufficiently large regular markets, in all random stable matchings, the matching type converges to the unique solution of the maximization problem (12). In particular, the fractions of badly matched individuals in all stable matchings converge to 0 and all stable matchings are *asymptotically proper*.

The hypothesis of Theorem 2 requires that either Assumption 2 holds and there exists a matching in which no individual stays single, or $j^* > 0$. Assumption 2 is always satisfied in the unrestricted roommate problem. In the marriage matching case, the Assumption holds as long as the number of men and women are equal. In such a case, the thesis of Theorem 2 holds even when $j^* = 0$. [12]If neither Assumption 2 nor $j^* > 0$ are satisfied, the thesis Theorem may not longer remain true.[13]

It follows from the proof of the Theorem, that the utilities of (asymptotically) all individuals converge to 1, i.e., to the top end of the support of the distribution of utilities. [14]

---

[12]In Menzel (2015), $J_N = N^{-1/2}$ and $j^* = 1$. The proofs in Menzel (2015) rely in multiple ways on the assumption that $j^* > 0$. Our proof demonstrates that this assumption is not necessary (at least, as long as Assumption 2 holds).

[13]Ashlagi, Leshno, and Kanoria (2016) show that in a marriage market with purely random preferences without any deterministic component (or, in our terminology, without any types), if the number of men and women are not equal and the size of the market grows indefinitely, the stable matching is essentially unique. Moreover, the utilities are essentially determined by the size of the imbalance in the market. We suspect that a similar phenomenon occurs in our case, with types.

[14]With large probability, for almost all i, the number of individuals who prefer i to the current match is of of order $N^{-1/2}$. Because i's match utility has to be higher than the utility from matching all but one of these individuals, i's utility will be close to 1, $1 - U_i \sim N^{-1/2}$.

One could interpret this observation as a statement about the efficiency of the stable matchings. At the same time, it is worth remembering that our model of stochastic utility is purely ordinal and the comparison of individual utilities through aggregate welfare might not always be justified. [15]

The convergence of the utilities to the top end of the support explains why only the derivatives at the top end play a role in the characterization of the matching. A similar phenomenon arises in Lee and Yariv (2014).

5.1. **Relation to Dagsvik (2000) and Menzel (2015).** The first-order conditions of problem (12) lead to the following system of equations[16]

$$\log d^* (t,t',p) = v^{t,t'} + \lambda_t + \lambda_{t'} \text{ for } (t,t') \notin R^*, \tag{13}$$

$$\log d^* (t,t,s) = v^t + \lambda_t,$$

---

[15]In cases where cardinal interpretation and the interpersonal comparison of utilities are warranted, the above observation confirms a recent result of Lee and Yariv (2014) who study a closely related random utility model, but with transferable utility.

[16]Consider first the non-restricted case. Because of (6d), we can take $d(t,t',p) = d(\{t,t'\},p)$. The Lagrangian associated with problem (12) can be written as

$$
\begin{aligned}
&\mathcal{L}(d,\lambda) \\
&= \sum_t d(t,t,s)\left(v^t - \log d(t,t,s)\right) + \frac{1}{2}\sum_t d(\{t\},p)\left(v^{t,t} - \log d(\{t\},p) - 1\right) \\
&\quad + \sum_{\{t,t'\}} d(\{t,t'\},p)\left(v^{t,t'} - \log d(t,t',p) - 1\right) \\
&\quad + \sum_t \lambda_t \left(d(t,t,s) + \sum_{t'} d(\{t,t'\},p)\right).
\end{aligned}
$$

The FOC are:

$$\log d(t,t,s) = -1 + \lambda_t + v^t,$$

$$\log d(\{t\},p) = -2 + 2\lambda_t + v^{t,t},$$

$$\log d(t,t',p) = -2 + \lambda_t + \lambda_{t'} + v^{t,t'} \text{ for } t \neq t'.$$

By changing the variable $\lambda_t \to \lambda_t - 1$, we obtain (13).

where λ_t are Lagrange multipliers associated with constraints (6a). In particular, the solution can be represented as a function of Lagrange multipliers λ_t. After substituting to the feasibility constraints, using the definitions (9), we obtain

$$\alpha_t e^{-\lambda_t} = V_t + \sum_s \alpha_s V^{t,s} \left(\alpha_s e^{-\lambda_s} \right)^{-1}.$$

Further, after letting $\Gamma_t = \alpha_t e^{-\lambda_t} - V_t$, we obtain

$$\Gamma_t = \sum_s \alpha_s \frac{V^{t,s}}{V_t + \Gamma_s}. \tag{14}$$

Equation (14) corresponds to equations (3.5) from Menzel (2015) (recall that in that paper, $j^* V^t = 1$.) Menzel (2015) refers to terms Γ_t as an inclusive value of type t individuals. Dagsvik (2000) and Menzel (2015) establish the uniqueness of the solution to (14) using the contraction properties of a certain operator. Here, we show that equation (14) describes a solution to the strictly concave maximization problem. Such a solution is, necessarily, unique.

5.2. **Proof of Theorem 2.** The probability that the matching type of *some* stable matching is ε-away from $d^* (\alpha, \upsilon)$ is bounded from above by

$$1 - \text{Prob} \left(\| d^\mu - d^* (\alpha, \upsilon) \| < \varepsilon \text{ for each stable matching } \mu \right) \tag{15}$$

$$= \text{Prob} \left(\| d^\mu - d^* (\alpha, \upsilon) \| > \varepsilon \text{ for some stable matching } \mu \right)$$

$$\leq \sum_{\mu, \| d^\mu - d^*(\alpha,\upsilon) \| > \varepsilon} \text{Prob} \left(\mu \text{ is a stable matching} \right)$$

$$= \sum_{d \in D_N(\alpha), \| d - d^*(\alpha,\upsilon) \| > \varepsilon} \sum_{\mu \in \mathcal{M}_N(d)} p(\mu)$$

$$= \sum_{d \in D_N(\alpha), \| d - d^*(\alpha,\upsilon) \| > \varepsilon} |\mathcal{M}_N(d)| \, p_N(d)$$

where we denote $p(\mu)$ as the probability that the realization of individual utilities makes matching μ stable, and

$$p_N(d) = \frac{1}{|\mathcal{M}_N(d)|} \sum_{\mu \in \mathcal{M}_N(d)} p(\mu)$$

is the average probability of a stable matching among matchings that induce d.

Below, we show the following claim: (15) converges to 0 at an exponential rate. Notice that

$$\frac{1}{N} \log \left(1 - \text{Prob} \left(\|d^\mu - d^* (\alpha, V)\| < \varepsilon \text{ for each stable matching } \mu\right)\right)$$

$$\leq \frac{1}{N} \log |D_N (\alpha)| + \max_{d \in D(\alpha), \|d - d^*(\alpha, V)\| > \varepsilon} \left(\frac{1}{N} \log |\mathcal{M}_N (d)| + \frac{1}{N} \log p_N (d)\right).$$

The number of type distributions, $|D_N (\alpha)|$, is polynomial in N.[17] Thus, $\frac{1}{N} \log |D_N (\alpha)| \to 0$ as N increases, and in order to prove our claim, it is enough to show that the second term of the last line above is strictly negative.

Let $j_N = N^{-1/2} J_N$, and let

$$\upsilon_N (d) = \sum_t d (t, t, s) (\log j_N V_t) + \frac{1}{2} \sum_{t,t'} d (t, t', p) \upsilon^{t,t'} + \sum_{t,t'} d (t, t', b) \upsilon^{t,t'}$$

be a version of the welfare where j^* is replaced by the approximating sequence $j_N \to j^*$. We have the following bound.

Lemma 2. *There exists a continuous function $g : R \to R_+$ such that $g(x) > 0$ for $x > 0$ and such that for each matching type $d \in D (\alpha, (\beta))$,*

$$\frac{1}{N} \log p_N (d) \leq \upsilon_N (d) - \max_{\delta \in D(\alpha, (\beta))} [\upsilon_N (\delta) + Entropy (\delta) + g (\delta (b))] \tag{16}$$

$$- \left(\frac{1}{2} d (p) + d (b)\right) \log N - \sum_t \alpha_t \log \alpha_t - \sum_t \beta_t \log \beta_t + o (1).$$

The reminder term $o(1)$ does not depend on type d.

The proof of Lemma 2 can be found in Appendix B. An overview of the main steps of the proof can be found in the section B.2 of the Appendix. Additionally, section 7 illustrates the main ideas behind the Lemma through back-of-envelope calculations in a special case of the model.

We can conclude the proof of the Theorem. Recall that Lemma 1 derives a bound on the number of matchings with type d. Together with Lemma 2, we have for each

---

[17]Notice that for each $d \in D_N (\alpha)$, $d(t, t',' c) \in \{0, \frac{1}{N}, ..., 1\}$ for each t, t', c. Thus, $|D_N (\alpha)| \leq N^{3|T|^2}$.

$d \in D\left(\alpha, (\beta_t)\right)$,

$$\frac{1}{N} \log |\mathcal{M}_N(d)| + \frac{1}{N} \log p_N(d)$$

$$\leq \upsilon_N(d) + \text{Entropy}(d) + g(d(b))$$

$$- \max_{\delta \in D(\alpha,(\beta))} [\upsilon_N(\delta) + \text{Entropy}(\delta) + g(\delta(b))] - g(d(b)) + o(1). \qquad (17)$$

As $N \to \infty$, the right hand side of inequality (17) is negative if either $d(b) > 0$, or $d(b) = 0$ and the value of the term in the first square bracket is strictly smaller than the value of the optimization problem

$$\max_{\delta \in D(\alpha,(\beta))} [\upsilon_N(\delta) + \text{Entropy}(\delta)].$$

If $j^* > 0$ or Assumption 2 holds, then the value of the above optimization problem converges to the value of (12). The result follows.

6. Noiseless limit

In this section, we consider a parametrized version of Example 1 with the match utility given by

$$U_i^j = \upsilon_{t_i}^{t_j} + \gamma \varepsilon_i^j,$$

where $\gamma > 0$ is a constant parameter, and ε_i^j are i.i.d. drawn from the distribution G that satisfies (11). As a normalization, we assume that $C_G = 1$. Such a model is essentially regular (see Example 1). By Theorem 2, as $N \to \infty$, the distribution over the matched types in all stable matchings converges to the unique solution

$$d^*\left(\alpha, \frac{1}{\gamma}\upsilon\right) = \arg \max_{d \in D(\alpha)} \text{Welfare}\left(d; \frac{1}{\gamma}\upsilon\right) + \text{Entropy}(d) \qquad (18)$$

$$= \arg \max_{d \in D(\alpha)} \text{Welfare}(d; \upsilon) + \gamma \text{Entropy}(d)$$

(Indeed, a monotonic transformation translates the model into its regular form with the "welfare" coefficients equal to $\frac{1}{\gamma}\upsilon_{\cdot}^{\cdot}$. The second line follows from the fact that function Welfare $(d; .)$ is linear in the "welfare" coefficients $\upsilon_{\cdot}^{\cdot}$.)

As the noise disappears, $\gamma \to 0$, the maximization problem (18) becomes dominated by the first (welfare) term, and the limit distribution over types in the stable matches

$v_t^{t'}$	$t' = H$	$t' = L$
$t = H$	1	0
$t = L$	0	-10

TABLE 1.

converges to the solution(s) of the welfare maximization problem

$$\max_{d \in D(\alpha)} \text{Welfare}(d; v). \tag{19}$$

This leads to a surprising discontinuity. As it is well-known, problem (19) characterizes outcomes that can be obtained in stable matches of the transferable utility (TU) model with no noise, $\gamma = 0$.[18] Typically, such outcomes do not arise as stable matchings in the non-transferable (NTU) utility model. As the noise converges to $\gamma \to 0$, the large population limit of the NTU random utility model converges to the TU outcome of the model with no noise, and not necessarily to the NTU outcome. [19]We illustrate this observation with a simple example.

Example 2. Suppose that the population is divided equally into two types, H and L. The "deterministic" utility v is described in Table 1. Let d^γ be the unique solution of (3), or the limit of type distributions in stable matchings. As $\gamma \to 0$, d^γ converges to the distribution d^* that assigns probability 1 to mixed pairs, $d^*(H, L) + d^*(L, H) = 1$. When $\gamma = 0$, d^* is the unique outcome of the TU model, and emphatically not the outcome of the NTU model. In the latter, the stable matching is assortative (i.e., it assigns a mass of zero to the mixed pairs).

In order to develop the intuition for this observation, it is useful to think about individuals as trading off the payoffs from the deterministic "utility" v and the random shocks ε. The H individuals are willing to match with an L individual if the random part of the utility is sufficiently high. On the other hand, because the difference in

---

[18]The classic result from Shapley and Shubik (1971) states that "any stable matching *in the marriage market* maximizes welfare" . It is easy to extend the observation to the roommate problem. The converse holds as well (see Chiappori, Galichon, and Salanie (2014)).

[19]The order of the limits is important. If we take the noise limit $\gamma \to 0$ first, and the large population limit $N \to \infty$ second, then the stable matchings trivially converge to the NTU case.

the deterministic values for L individuals is much higher, the latter are willing to give up on the random shock in order to secure the larger deterministic value. In a sense, the random shocks in the random non-transferable utility model play a role of transfers. This observation can be interpreted as a non-transferable foundation for the transferable utility model.

7. Example: No types

We are going to illustrate some of the ideas of the proofs on a simple example. We consider a setting in which there are no (non-trivial) types, or that $T = \{*\}$. We assume that there are no match-specific shocks, all the match utilities are drawn from the uniform distribution on the interval $[0, 1]$, and the utility of being single is drawn i.i.d. from distribution

$$P\left(U_i^i \leq 1 - x\right) = (1 - x)^{J_N} \tag{20}$$

(as maximum of J_N random draws of uniformly distributed variables). In such a case, a type distribution d of the matching can be characterized by a triple $(d(s), d(p), d(b))$ of positive numbers such that $d(s) + d(p) + d(b) = 1$, i.e., the fractions of single, properly and badly matched agents.

Even in this very special case, our result is non-trivial and it establishes that , as $N \to \infty$, the fraction of badly matched agents $d(b)$ in any stable (possibly, improper) matching converges to 0. The result follows immediately from Lemmas 1, 2, and the discussion of the proof of Theorem 2 in Section 5.2. Below, we heuristically describe the proofs of the two key lemmas.

7.1. **Number of matchings.** First, we derive an upper bound on the number of matchings with a type distribution d. We do it in four steps. First, we find the number of ways of dividing the population into single, properly and badly matched individuals. This number is equal to

$$\frac{N!}{(Nd(s))!\,(Nd(p))!\,(Nd(b))!}. \tag{21}$$

Second, we find the number of proper matchings among $Nd(p)$ properly matched individuals. Each such a matching can be uniquely created in the following procedure. Fix an arbitrary ranking of all individuals. In each step of the procedure, pick an

individual with the highest remaining rank and choose an arbitrary match for her. Repeat the step until there are no unmatched individuals. There are $(Nd(p) - 2k + 1)$ different ways of choosing a match in the kth step of the procedure. Thus, the number of different proper matchings is equal to

$$(Nd(p) - 1) \cdot (Nd(p) - 3) \dots \cdot 3 \cdot 1 \leq ((Nd(p))!)^{1/2}. \tag{22}$$

Third, we find an upper bound on the number of matchings among $Nd(b)$ individuals, where all individuals are badly matched. It is enough to find a number of all improper matchings. Each improper matching is a bijection from the set of $Nd(b)$ individuals onto itself. The number of bijections is equal to[20]

$$(Nd(b))!. \tag{23}$$

Finally, a bound on the number of matchings with type d, $|\mathcal{M}_N(d)|$, can be obtained by multiplying (21)-(23). Using Stirling's approximation[21], we obtain

$$\frac{1}{N} \log |\mathcal{M}_N(d)|$$
$$\leq \log N - 1 - d(s) \log(N(ds)) + d(s) - \frac{1}{2}d(p) \log(Nd(p)) + \frac{1}{2}d(p)$$
$$= \left(d(b) + \frac{1}{2}d(b)\right) \log N - \left(d(b) + \frac{1}{2}d(b)\right) - d(s) \log d(s) - \frac{1}{2}d(p) \log d(p),$$

which corresponds to Lemma 1.

7.2. **Probability of stable matching.** Next, we illustrate the ideas behind Lemma 2. We are going to derive an upper bound on the (average) probability that a matching with type d is stable. The argument proceeds in fours steps.

In the first step (Lemma 4 in the Appendix), we derive a bound on the probability that a particular matching μ is stable. The key idea is to notice that, conditionally on the realization of the match utilities, all the other relevant events are independent.

---

[20]To see it, imagine that each of the $Nd(b)$ individuals has two copies, "left" and "right". Each (improper) matching μ can be uniquely identified with a directed graph with an arrow pointing from the "left" copy of each individual i to the "right" copy of each individual $\mu(i)$. Each "left" copy has exactly one arrow leaving out and each "right" copy has exactly one arrow pointing towards it. The number of graphs so constructed is equal to (23).

[21]$\log N! \approx N \log N - N$.

Let $U_i^{\mu(i)}$ be the realized utility of individual i in her match with $\mu(i)$. It is convenient to define $x_i = 1 - U_i^{\mu(i)}$. Because the utilites are drawn from the uniform distribution, x_i is the probability that an independently drawn utility is higher than the match utility of i. Also, let

$$x = \frac{1}{N} \sum_i x_i. \tag{24}$$

We begin with the conditional probability that μ is stable given the realization of x_is. Matching μ is stable, if the following events are simultaneously satisfied:

- for all i that are not single, the utility of staying single must be lower than the match utility. The conditional probability of such an event is equal to

$$\prod_{i:\mu(i)\neq i} (1 - x_i)^{J_N} \leq \prod_{i:\mu(i)\neq i} (1 - x_i)^{J_N - 1}, \tag{25}$$

- no pair $\{i, j\} \in S(\mu)$, where $S(\mu) = \{\{i, \mu(i)\} : i \in I\}$ is the set of matched pairs, blocks the matching. Because all the blocking pairs are independent, the probability that such an event hold can be bounded by

$$\prod_{\{i,j\}\notin S(\mu):i\neq j} (1 - x_i x_j), \tag{26}$$

- inequality (5) holds for each i. Because of independence, the probability can be bounded by

$$\prod_{i:\mu(i)\neq \mu^{-1}(i)} x_i. \tag{27}$$

All the above events are independent, and the product of (25)-(27) leads to an upper bound on the conditional probability that μ is stable. A bound on the unconditional probability $p(\mu)$ can be obtained by integrating the conditional bound over the realization of variables x_i:

$$p(\mu) \leq J_N^{d(s)N} \iint \left(\prod_{i\{i,j\}\notin S(\mu):i\neq j} (1 - x_i x_j) \prod_{i:\mu(i)\neq \mu^{-1}(i)} x_i \right) \left(\prod_i (1 - x_i)^{J_N - 1} \right) \prod_i dx_i.$$

(Observe that $(1 - x_i)^{J_N - 1}$ terms come from (25) if i is matched, and, if i is single, from the fact that the density of the utility of being single (20) is equal to

$$J_N (1 - x)^{J_N - 1} dx.)$$

The purpose of the second step (Lemma 5) is to simplify the above expression. Because $1 - a \leq e^{-a}$ for each a, we have

$$\prod_{\{i,j\} \notin S(\mu):i \neq j} (1 - x_i x_j) \leq e^{-\sum_{i\{i,j\} \notin S(\mu):i \neq j} x_i x_j} \approx e^{-\frac{1}{2} \sum_{i,j} x_i x_j} = e^{-\frac{1}{2}(Nx)^2} = e^{-\frac{1}{2}N^2 x^2},$$

$$\prod_i (1 - x_i)^{J_N - 1} \leq e^{-(J_N - 1)Nx} \approx e^{-J_N Nx} = e^{-N^{3/2} j_N x}.$$

(The approximations are correct because the set of pairs that are excluded from the summation is small relative to all the other pairs.) Further, Lemma 8 shows that the average of expression (27) over all matchings with type d is not larger than

$$\frac{1}{|\mathcal{M}_N(d)|} \sum_{\mu:d^\mu = d} \left(\prod_{i:\mu(i) \neq \mu^{-1}(i)} x_i \right) \leq x^{Nd(b)} = e^{Nd(b) \log x}.$$

Together, these two observations imply that the average probability that a matching with type d is stable can be bounded by

$$p_N(d) \leq e^{Nd(s)\left(\log j_N + \frac{1}{2} \log N\right)} \int e^{-\frac{1}{2}N^2 x^2 - N^{3/2} j_N x + Nd(b) \log x} d\mu^N(x).$$

Here, μ^N is a probability distribution of the average (24) of N uniformly distributed variables.

In the third step (Lemma 6), we rely on the large deviation theory to show that, for our purposes, μ^N is well-approximated by[22]

$$d\mu^N(x) \approx e^{N(1 + \log x)} dx,$$

which implies

$$p_N(d) \leq e^{Nd(s)\left(\log j_N + \frac{1}{2} \log N\right)} \int e^{-\frac{1}{2}N^2 x^2 - N^{3/2} j_N x + Nd(b) \log x + N \log x + N} d(x).$$

---

[22]The approximation follows from the Chernoff bound applies and

$$\mu^N \left(\frac{1}{N} \sum_i x_i < x \right) \leq \mu^N \left(e^{-\sum_i x^i/x} \geq e^{-N} \right) \leq e^N E_{\lambda^N} e^{-\sum_i x^i/x} = e^N E_{\lambda^N} \left(e^{-\sum_i x^i/x} \right) \leq e^N x^N = e^{N(1 + \log x)}.$$

This roughly yields an approximation $d\mu^N(x) = Ne^N x^{N-1} = e^{N(1 + \log x)} \left(N\frac{1}{x} \right) \approx e^{N(1 + \log x)}$ (the term in brackets does not affect the convergence properties of the integral). The precise argument is given in Lemma 11.

The convergence properties of the integral can be analyzed by maximizing the expression in the exponent: for large N

$$\frac{1}{N} \log p_N(d) \le d(s) \left(\log j_N + \frac{1}{2} \log N \right) + \max_{x \in [0,1]} \left[-\frac{1}{2} N x^2 - j_N N^{1/2} x + (1 + d(b)) \log x + 1 \right]$$

$$\le \upsilon_N(d) - \left(\frac{1}{2} d(p) + d(b) \right) \log N + 1 + \max_{y \ge 0} \left[-\frac{1}{2} y^2 - j_N y + (1 + d(b)) \log y \right],$$

$$(28)$$

where the second inequality came from substitution $y = N^{1/2} x$, and the fact that $1 + d(b) = d(s) + d(p) + 2d(b)$.

In the last step (Lemma 7) we bound the value of the maximization problem on the right-hand side of the above inequality. Take any type distribution δ such that $\delta(b) = d(b)$. Then,

$$-\frac{1}{2} y^2 - j_N y + (1 + d(b)) \log y$$

$$= -\frac{1}{2} y^2 - j_N y + \left(\left(\frac{1}{2} \delta(p) + \delta(b) \right) \log y^2 + \delta(s) \log y \right)$$

$$\le \left(-\frac{1}{2} y^2 + \left(\frac{1}{2} \delta(p) + \delta(b) \right) \log y^2 \right) + \left(-j_N y + \delta(s) \log y \right).$$

Because $-ax + \log x \le -\log a - 1$ for each $a.x \ge 0$, the above is not larger than

$$\le \frac{1}{2} (\delta(p) + 2\delta(b)) \log (\delta(p) + 2\delta(b)) + \delta(s) \log \delta(s) - \delta(s) \log j_N - \left(\delta(s) + \frac{1}{2} \delta(p) + \delta(b) \right)$$

$$= \frac{1}{2} (\delta(p) + 2\delta(b)) \log (\delta(p) + 2\delta(b)) + \delta(s) \log \delta(s) - \delta(s) \log j_N - 1 + \frac{1}{2} \delta(p).$$

Further, straightforward arguments show that the above is not larger (and strictly smaller if $\delta(b) > 0$) than[23]

$$\le \frac{1}{2} \delta(p) \log \delta(p) + \delta(b) + \delta(s) \log \delta(s) - \delta(s) \log j_N - 1 + \frac{1}{2} \delta(p)$$

$$= -\upsilon_N(\delta) - \text{Entropy}(\delta) - \delta(b) \log \delta(b) - 1.$$

Together with (28), we obtain

$$\frac{1}{N} \log p_N(d) \le \upsilon_N(d) - \left(\frac{1}{2} d(p) + d(b) \right) \log N - [\upsilon_N(\delta) + \text{Entropy}(\delta)] - \delta(b) \log \delta(b),$$

---

[23]Notice that $\delta(b) - \frac{1}{2} (\delta(p) + 2\delta(b)) \log (\delta(p) + 2\delta(b))$ is increasing and strictly concave in $\delta(b)$ for $\delta(p), \delta(b) \in [0, 1]$.

which corresponds to the inequality from the statement of Lemma 2.

8. Conclusions

In this paper, we consider the large random utility matching model introduced in Dagsvik (2000). We extend the result from Menzel (2015) to the roommate problem. Since the Gale-Shapley algorithm does not apply to the roommate problem, our proof technique is necessarily very different. Instead, our proof is based on the estimates of the number of different matchings and an estimates of the probability for each matching being stable. The proof leads to a succinct characterization of type distributions obtained in stable matchings as maximizers of a "welfare plus entropy" formula. The formula leads to a surprising connection between a noiseless limit of the non-transferable model with large population.

It is natural to ask whether the model and convergence results can be extended further, for instance, to many-to-many matching. One difficulty is that our results rely on Tan (1991)'s characterization of approximate stable matching in the roommate problem. As far as we know, there are no appropriately tight approximate solutions concepts for the general many-to-many matching models that can be used in the probability estimates as in the Lemma 2. (Some of the most recent results on the existence of approximate solutions can be found in Che, Kim, and Kojima (2015).)

An alternative approach is to abandon classic matching theory and analyze many-to-many matching in the game-theoretic framework of search theory. The advantage of such approach is that the existence of equilibrium is assured by standard arguments. We follow this approach in companion paper Peski (2015), where we consider a search-based model of many-to-many matching with the utility extending (2). We show that the equilibrium distribution over the match types can be characterized as critical points of a "welfare plus entropy" formula. In the special case of the roommate problem with one-to-one matching, the characterizations from this paper and from Peski (2015) coincide.

Appendix A. Number of matchings: proof of Lemma 1

For each (improper) matching μ, define sets

$$A^\mu\left(t, t', c\right) = \left\{i : t_i = t, t_{\mu(i)} = t', c^\mu\left(i\right) = c\right\}.$$

Sets $A^\mu\left(.,.,.\right)$ are disjoint and $\left|A^\mu\left(t, t', c\right)\right| = N d^\mu\left(t, t', c\right)$. Set $A^\mu\left(t, t', c\right)$ consists of type t individuals who are matched with type t' individuals and the category of their match is c.

From now on, we fix the type $d \in D\left(\alpha, \left(\beta_t\right)\right)$. To shorten the notation, for each $t, t' \in T$ and $c \in \{s, p, b\}$, we write $d\left(c\right)$ instead of $\sum_{t, t'} d\left(t, t', c\right)$, $d\left(t, ., c\right)$ instead of $\sum_{t'} d\left(t, t', c\right)$ and $d\left(., t', c\right)$ instead of $\sum_t d\left(t, t', c\right)$.

The counting of matchings with type d is divided into the following steps.

Different choices of sets A^μ. The number of different ways of dividing the population into sets $A^\mu\left(t, t', c\right)$ is equal to

$$\prod_t \left(\frac{\left(\alpha_t N\right)!}{\prod_{t', c}\left(N d\left(t, t', c\right)\right)!}\right). \tag{29a}$$

Proper matchings between individuals in sets $A^\mu\left(t, t', p\right)$ and $A^\mu\left(t', t, p\right)$ for $t \neq t'$. Any two pairwise matchings between two disjoint sets with the same cardinality can be obtained from each other by the permutation of the elements of one of the sets. For this reason, the number of pairwise matchings between $A^\mu\left(t, t', p\right)$ and $A^\mu\left(t', t, p\right)$ is equal to

$$\left(N d\left(t, t', p\right)\right)!.$$

The number of proper matchings that can be formed by the individuals in set $\bigcup_{t \neq t'} A^\mu\left(t, t', p\right)$ so that individuals in $A^\mu\left(t, t', p\right)$ match with $A^\mu\left(t', t, p\right)$ is equal to

$$\prod_{\{t, t'\}: t \neq t'}\left(N d\left(t, t', p\right)\right)! = \left(\prod_{(t, t'): t \neq t'}\left(N d\left(t, t', p\right)\right)!\right)^{1/2}. \tag{29b}$$

The square root on the right-hand side comes from the fact that each two-element subset in the product on the left-hand side corresponds to exactly two ordered pairs on the right-hand side.

Proper matchings between individuals in sets $A^\mu\left(t, t, p\right)$. If set $A^\mu\left(t, t, p\right)$ has an odd number of elements, then there are no pairwise matchings with set $A^\mu\left(t, t, p\right)$.

Otherwise, for each t, the number of pairwise matchings within population of type t individuals is not larger to the number of ways of choosing $\frac{1}{2}Nd(t,t,p)$ indistinguishable pairs from $Nd(t,t,p)$-element sets, or

$$\frac{((Nd(t,t,p))!)}{2^{\frac{1}{2}Nd(t,t,p)}\left(\left(\frac{1}{2}Nd(t,t,p)\right)!\right)} = 1 \cdot 3 \cdot ... \cdot (Nd(t,t,p)-1) \le ((Nd(t,t,p))!)^{1/2}.$$

The number of proper matchings that can be formed by the individuals in set $\bigcup_t A^\mu(t,t,p)$ so that each type t individual is matched with another type t is not larger than

$$\left(\prod_t (Nd(t,t,p))!\right)^{1/2}. \tag{29c}$$

Bad matchings. In order to count the number of bad matchings, notice that each bad matching can be uniquely represented as a directed graph with the set of nodes equal to the sets of two copies, "left" and "right", of badly matched individuals, and the edges connecting a "left" copy of i to the "right" copy of j if and only if $j = \mu(i)$. Because μ is a bijection, each "left" node has exactly one outgoing edge, and each "right" node has exactly one incoming edge.

Each "left" copy of an individual in set $A^\mu(t,t',b)$ is connected to a "right" node of type t' individual that belongs to set

$$B^\mu(t',t,b) = \left\{j : t_j = t', t_{\mu^{-1}(j)} = t, c^\mu(j) = b\right\}.$$

The number of individuals in set $B^\mu(t',t,b)$ is equal to $d^\mu(t,t',b)N$. The number of ways of dividing the set of "right" nodes of type t' individuals into sets $B^\mu(t',t,b)$ is equal to

$$\prod_{t'} \frac{(Nd(.,t',b))!}{\prod_t (Nd(t,t',b))!}.$$

Given the definition of sets $B^\mu(.)$, each directed graph corresponds to a (proper) matching between disjoint sets given that a "left" node in set $A^\mu(t,t',b)$ must be matched with a "right" node in set $B^\mu(t',t,b)$. The number of such directed graphs is equal to

$$\prod_{t,t'} (Nd(t,t',b))!.$$

Thus, the total number of bad matchings between all badly matched individuals is equal to

$$\prod_{t'} \frac{(Nd\,(.,t',b))!}{\prod_t (Nd\,(t,t',b))!} \prod_{t,t'} (Nd\,(t,t',b))! = \prod_{t'} (Nd\,(.,t',b))! = \prod_t (N\beta_t\,(d))! \qquad (29\text{d})$$

The last equality follows from (6e).

All matchings. The number of all improper matchings with type d is not larger than the product of (29a)-(29d):

$$|\mathcal{M}\,(d)|$$

$$\leq \prod_t \left(\frac{(\alpha_t N)!}{\prod_{t',c} (Nd\,(t,t',c))!} \right) \left(\prod_{t,t'} (Nd\,(t,t',p))! \right)^{1/2} \prod_t (N\beta_t\,(d))!$$

$$= \prod_t ((\alpha_t N)!\,(N\beta_t\,(d))!) \left(\frac{1}{\prod_t (Nd\,(t,t,s))!} \right) \left(\frac{1}{\prod_{t,t'} (Nd\,(t,t',p))!} \right)^{1/2} \left(\prod_{t,t'} \frac{1}{(Nd\,(t,t',b))!} \right).$$

(We use the fact that $d\,(t,t',s) = 0$ for $t \neq t'$ and the convention that $0! = 1$.) The Stirling approximation implies that for each $a \in [0,1]$,

$$\frac{1}{N} \log\,(aN)! = a \log a + a\,(\log N - 1) + o\,(1),$$

where $o\,(1)$ does not depend on $a \in [0,1]$. Thus,

$$\frac{1}{N} \log |\mathcal{M}\,(d)|$$

$$\leq \sum_t \alpha_t \log \alpha_t + \log N - 1$$

$$+ \sum_t \beta_t\,(d) \log \beta_t\,(d) + d\,(b)\,(\log N - 1)$$

$$- \sum_t d\,(t,t,s) \log d\,(t,t,s) - d\,(s)\,(\log N - 1)$$

$$- \frac{1}{2} \sum_{t,t'} d\,(t,t',p) \log d\,(t,t',p) - \frac{1}{2} d\,(p)\,(\log N - 1)$$

$$- \sum_{t,t'} d\,(t,t',b) \log d\,(t,t',b) - d\,(b)\,(\log N - 1) + o\,(1).$$

The result follows from the definition of entropy (7) and the fact that $d\,(s) + d\,(p) + d\,(b) = 1$.

APPENDIX B. PROOF OF LEMMA 2

B.1. **Preliminary observation.** Let $\mathcal{F} = \left\{ F_t^{t',\theta}, F_t \right\}_{t,t',\theta}$ be a regular random utility model. For each $F \in \mathcal{F}$, let $V^F = \frac{dF}{dU}(1)$ denote the associated "welfare" constant. We use the following observation about regular model $\mathcal{F}$:

Lemma 3. *There exist continuous, strictly increasing and continuously differentiable functions $H, h : [0,1] \to R_+$ such that $h(1) \leq 1$,*

$$h(0) = H(0) = 0,$$

$$h'(0) = H'(0) = 1,$$

$$h(x) \leq x \leq H(x) \text{ for each } x,$$

and, for each $F \in \mathcal{F}$,

$$V^F h(1-u) \leq 1 - F(u) \leq V^F H(1-u), \text{ and } \frac{dF}{dU}(u) \leq V^H H'(1-u).$$

Proof. Let $\mathcal{G}$ be a collection of functions $G(x) = \frac{1}{V^F}(1 - F(1-x))$ for each $F \in \mathcal{F}$. Each $G \in \mathcal{G}$ is continuous, continuously differentiable, $G(0) = 0$, and $G'(0) = 1$. Because set $\mathcal{G}$ is finite, it follows that there exists functions functions $H, h : [0,1] \to R_+$ that have all the required properties stated in the Lemma and such that for each $G \in \mathcal{G}$,

$$h(x) \leq G(x) \leq H(x), \text{ and } G'(x) \leq H'(x).$$

$\square$

Let $H^* = \max_{x \in [0,1]} \left(\frac{1}{h'(x)}, H'(x) \right) < \infty$. The constant is bounded due to the fact that h is continuously differentiable on a compact interval. Let $V^* = \max_F \left\{ V_t, V_t^{t',\theta} \right\}$.

From now on, fix a matching type $d \in D(\alpha, (\beta_t))$. Let $\beta = \sum_t \beta_t$. Let $S_2 = \{\{i,j\} : i,j \in I, i \neq j, (i,j) \notin R\}$ be the set of all (feasible) match pairs and let $S = S_2 \cup I$ be the set of all possible matches, that includes match pairs as well as the possibility of staying alone. Let $S(\mu) = \{\{i, \mu(i)\} : i \in I\}$ be the set of matches that occur in a matching μ. Let $B(\mu) = \{i : \mu(\mu(i)) \neq i\}$ be the set of badly matched individuals.

B.2. **Overview of the proof.** We describe the main steps of the proof, leaving details for subsequent sections.

In the first step, we derive a bound on the probability $p(\mu)$ that a given matching μ is stable. The idea is to consider first the conditional probability that μ is stable, conditionally on the realization of the match utilities $U_i^{\mu(i)}$ as well as the match specific shocks $\theta^{\{i,j\}}$. Such a probability is a product of the probabilities of independent events:

- for each i such that $\mu(i) \neq i$, the event that J^N i.i.d. draws from distribution F_{t_i} is not larger than $U_i^{\mu(i)}$,
- for each unmatched pair of individuals $\{i, j\} \in S_2 \setminus S(\mu)$, the complement of the event that both U_i^j is larger than $U_i^{\mu(i)}$ and U_j^i is larger than $U_j^{\mu(i)}$,
- for each badly matched individual $i \in B(\mu)$, the event that the utility from the "better" match $U_i^{\mu^{-1}(i)}$ is larger than the utility from the worse match $U_i^{\mu(i)}$.

Using approximations from Lemma 3, integrating over the normalized utilities and match specific shocks, and changing the variables

$$x_i = h\left(1 - U_i^{\mu(i)}\right) \approx 1 - U_i^{\mu(i)}$$

(where the approximation is good for utility values close to the top end of the support $U_i^{\mu(i)} \sim 1$), we obtain the following bound:

Lemma 4. *There exists a continuous function $a(x)$ such that $a(0) = 0$ and for each N, each matching μ,*

$$p(\mu) \leq J_N^{d(s)N} \left(\prod_{\{i,j\} \in S(\mu)} V^{t_i, t_j} \right) \underset{0 \leq x_i \leq 1}{\int \cdots \int} \left[\prod_{i \in B(\mu)} x_i \prod_i (1 + a(x_i)) \ldots \right. \tag{30}$$

$$\left. \ldots \prod_i (1 - V_{t_i} x_i)_+^{J_N - 1} \prod_{\{i,j\} \in S_2 \setminus S(\mu)} \left(1 - V^{t_i, t_j} x_i x_j\right)_+ \right] \prod_i dx_i.$$

(We write $(x)_+ = \max(x, 0)$.) The expression in the square bracket can be treated as a function of N independent and uniformly distributed variables X_i and the integral is the expectation of such a function.

In the second step, we find a bound on the average probability $p_N(d)$ of a stable matching with type d. The bound has the form of an integral of an expression that

depends on the *average* of (normalized) utilities across the population. For each $x : I \to [0, 1]$, each type t, let

$$x_t = \frac{1}{\alpha_t N} \sum_{i:t_i=t} x_i.$$

Let μ^m be the probability distribution of the average $\frac{1}{m} \sum_{i=1}^m X_i$, where X_i are independent and uniformly distributed on the interval $[0, 1]$. We use bound (30) , an inequality,

$$\prod_i (1 - x_i)_+ \le e^{-\sum_i x_i}$$

and some algebra to show

Lemma 5. *There exists a continuous function $A(x)$ such that $A(0) = 0$ and for each N, each matching type d*

$$p_N(d) \tag{31}$$

$$\le e^{N\left(v_N(d) + \frac{1}{2}d(s)\log N\right)}$$

$$\cdot \int\cdots\int_{0 \le x_t \le 1} \exp\left(-N^{3/2} j_N \sum_t V_t \alpha_t x_t - \frac{1}{2} N^2 \sum_{(t,t')\in R^*} V^{t,t'} \alpha_t x_t \alpha_{t'} x_{t'}\right)$$

$$\cdot \exp\left(N \sum_t \beta_t \log x_t + N A\left(\sum_t \alpha_t x_t\right)\right) \prod_t d\mu^{\alpha_t N}(x_t)$$

The goal of the third step is to estimate the convergence rate of the integral (31) as $N \to \infty$. We rely on two ideas from the large deviation theory. First, we use the fact that for large m, the following approximation of measure μ^m can be used for $x \sim 0$:

$$d\mu^m(x) \approx e^{m(1+\log x)}.$$

Second, Varadhan's Lemma (Ellis (2005)) says that for large m, the value of the integral grows at the rate of the largest exponent of the integrand:

$$\frac{1}{m} \log \int e^{mF(x)} dx \to \max_x F(x).$$

Using (a careful version of) the above approximations, we can bound the convergence rate of integral (31) with the solution to a certain maximization problem:

Lemma 6. *When $N \to \infty$,*

$$\frac{1}{N} \log p_N(d) \leq \upsilon_N(d) + \frac{1}{2} d(s) \log N + 1 + o(1) \tag{32}$$

$$+ \max_{x. \in [0,1]^T} \left\{ \begin{array}{c} -N^{1/2} j_N \sum_t V_t \alpha_t x_t - \frac{1}{2} N \sum_{(t,t') \in R^*} V^{t,t'} \alpha_t x_t \alpha_{t'} x_{t'} \\ + \sum_t (\alpha_t + \beta_t) \log x_t + A\left(\sum_t \alpha_t x_t\right) \end{array} \right\}.$$

In the fourth step, we find an upper bound on the value of the maximization problem in (32).

Lemma 7. *There exists a continuous function $g : R \to R_+$ such that $g(x) > 0$ for $x > 0$ and such that as $N \to \infty$,*

$$\limsup_{N \to \infty} \max_{x. \in [0,1]^T} \left\{ \begin{array}{c} -N^{1/2} j_N \sum_t V_t \alpha_t x_t - \frac{1}{2} N \sum_{(t,t') \in R^*} V^{t,t'} \alpha_t x_t \alpha_{t'} x_{t'} \\ + \sum_t (\alpha_t + \beta_t) \log x_t + A\left(\sum_t \alpha_t x_t\right) \end{array} \right\}$$

$$\leq -1 - \sum_t \alpha_t \log \alpha_t - \sum_t \beta_t \log \beta_t - \left(\frac{1}{2} d(s) + \frac{1}{2} d(p) + d(b)\right) \log N$$

$$- \max_{\delta \in D(\alpha, (\beta))} [\upsilon_N(\delta) + Entropy(\delta) + g(\delta(b))] + o(1).$$

Lemma 2 follows from Lemmas 6 and 7

B.3. **Proof of Lemma 4: Probability that matching μ is stable.** Let $\tilde{\Theta} : S \to \Theta$ be a realization of match-specific shocks. Similarly, let $\tilde{U} : I \to R$ denote the realizations of utilities $U_i^{\mu(i)}$ of agents in their (worse) matches. When the values of $\tilde{\Theta}$ are clear from the context, we write $\tilde{F}_i^j$ instead of $F_{t_i}^{t_j, \tilde{\Theta}\{i,j\}}$, and $\tilde{V}_i^j$ instead of $V_{t_i}^{t_j, \tilde{\Theta}\{i,j\}}$ for each i, j.

B.3.1. *Conditional probability given $\tilde{U}$ and $\tilde{\Theta}$.* First, we estimate the conditional probability that μ is a stable matching given the realization of match-specific $\tilde{\Theta}$ and utilities $\tilde{U}$. By conditional independence assumptions, the conditional probability $p\left(\mu|\tilde{\Theta}, \tilde{U}\right)$ is equal to the product of

- the probability that the individual rationality is preserved for all individuals who are alone:

$$\prod_{i:\mu(i) \neq i} \left(F_{t_i}\left(\tilde{U}_i\right)\right)^{J_N},$$

- the probability that no unmatched pair $\{i, j\}$ blocks an existing match:

$$\prod_{\{i,j\}\in S_2\backslash S(\mu)} \left(1 - \left(1 - \tilde{F}_i^j\left(\tilde{U}_i\right)\right)\left(1 - \tilde{F}_j^i\left(\tilde{U}_j\right)\right)\right), \text{ and}$$

- the probability that the utility of badly matched individuals in their "better" match is not smaller than the utility in the "worse" match:

$$\prod_{i\in B(\mu)} \left(1 - \tilde{F}_i^{\mu^{-1}(i)}\left(\tilde{U}_i\right)\right).$$

B.3.2. *Conditional probability given* $\tilde{\Theta}$. The conditional probability $p\left(\mu|\tilde{\Theta}\right)$ that μ is stable given $\tilde{\Theta}$ is equal to the above product integrated over the realization of utilities $\tilde{U}$. Conditionally on the realization of match-specific shock, the variables $\tilde{U}_i$ are independently distributed with density

$$d\tilde{F}_i^j\left(\tilde{U}_i\right) \text{ if } \mu\left(i\right) \neq i, \text{ and}$$

$$J_N\left(F_{t_i}\left(\tilde{U}_i\right)\right)^{J_N-1} dF_{t_i}\left(\tilde{U}_i\right) \text{ if } \mu\left(i\right) = i.$$

Thus, the conditional probability $p\left(\mu|\tilde{\Theta}\right)$ that μ is stable given $\tilde{\Theta}$ is equal to

$$p\left(\mu|\tilde{\Theta}\right) = \int\cdots\int_{\tilde{U}_i, i\in I} \prod_{i:\mu(i)\neq i} \left(F_{t_i}\left(\tilde{U}_i\right)\right)^{J_N} \left(\prod_{i\in B(\mu)} \left(1 - \tilde{F}_i^{\mu^{-1}(i)}\left(\tilde{U}_i\right)\right)\right)$$

$$\prod_{\{i,j\}\in S_2\backslash S(\mu)} \left(1 - \left(1 - \tilde{F}_i^j\left(\tilde{U}_i\right)\right)\left(1 - \tilde{F}_j^i\left(\tilde{U}_j\right)\right)\right) \tag{33}$$

$$\prod_{i:\mu(i)=i} J_N\left(F_{t_i}\left(\tilde{U}_i\right)\right)^{J_N-1} dF_{t_i}\left(\tilde{U}_i\right) \prod_{i:\mu(i)\neq i} d\tilde{F}_i^{\mu(i)}\left(\tilde{U}_i\right).$$

B.3.3. *Unconditional probability*. Let $y_i = 1 - \tilde{U}_i$. Lemma 3, the fact that $F_{t_i}\left(\tilde{U}_i\right) \leq 1$, and equation (33) imply

$$p\left(\mu; \tilde{\Theta}\right) \leq J_N^{d(s)N} \int\cdots\int_{\tilde{U}_i, i\in I} \prod_i \left(1 - V_{t_i} h\left(y_i\right)\right)^{J_N-1} \prod_{\{i,j\}\in S_2\backslash S(\mu)} \left(1 - \tilde{V}_i^j \tilde{V}_j^i h\left(y_i\right) h\left(y_j\right)\right)$$

$$\left(\prod_{i\in B(\mu)} \tilde{V}_i^{\mu^{-1}(i)} H\left(y_i\right)\right) \prod_{i:\mu(i)=i} V_{t_i} \prod_{i:\mu(i)\neq i} \tilde{V}_i^{\mu(i)} \prod_i H'\left(y_i\right) dy_i.$$

Because the match shocks are independent, and $V^{t_i,t_j} = \sum_\theta \Phi^{t_i,t_j}(\theta)\tilde{V}_i^j \tilde{V}_j^j$, we obtain

$$p(\mu) \leq J_N^{d(s)N} \left(\prod_{i:\mu(i)=i} V_{t_i} \prod_{\{i,j\} \in S_2 \cap S(\mu)} V^{t_i,t_j} \right)$$

$$\int \cdots \int_{0 \leq y_i \leq 1, i \in I} \prod_i (1 - V_{t_i} h(y_i))^{J_N-1} \prod_{\{i,j\} \in S_2 \backslash S(\mu)} \left(1 - V^{t_i,t_j} h(y_i) h(y_j)\right)$$

$$\left(\prod_{i \in B(\mu)} H(y_i) \right) \prod_i H'(y_i)\, dy_i.$$

By another change of variables $x_i = h(y_i)$, the value of the integral from the above expression is not larger than

$$\leq \int \cdots \int_{0 \leq x_i \leq h(1), i \in I} \prod_i (1 - V_{t_i} x_i)^{J_N-1} \prod_{\{i,j\} \in S_2 \backslash S(\mu)} \left(1 - V^{t_i,t_j} x_i x_j\right) \left(\prod_{i \in B(\mu)} x_i \right) \qquad (34)$$

$$\prod_i \max\left(1, \frac{H(h^{-1}(x_i))}{x_i}\right) \prod_i \left(\frac{H'(h^{-1}(x_i))}{h'(h^{-1}(x_i))} \right) dx_i.$$

Let

$$a(x) = \left(\max\left(1, \frac{H(h^{-1}(x))}{x}\right) \right) \frac{H'(h^{-1}(x))}{h'(h^{-1}(x))} - 1 \text{ for } x > 0,$$

and $a(0) = 0$. By properties of functions h and H listed in Lemma 3, $a(x)$ is continuous. Lemma 4 follows from the fact that $h(1) \leq 1$.

B.4. **Proof of Lemma 5: Average probability that a matching of type d is stable.** The proof follows from (30) and the subsequent observations. First, observe that

$$\frac{1}{N} \log J_N^{d(s)N} \left(\prod_{i:\mu(i)=i} V_{t_i} \prod_{\{i,j\} \in S_2 \cap S(\mu)} V^{t_i,t_j} \right)$$

$$= d(s) \log J_N + \frac{1}{N} \left(\sum_{i \in S(\mu) \backslash S_2} \log V_{t_i} + \sum_{\{i,j\} \in S(\mu) \cap S_2} \log V^{t_i,t_j} \right)$$

$$= d(s) \left(\log j_N + \frac{1}{2} \log N \right) + \sum_t d(t,t,s) \log V_t + \frac{1}{2} \sum_{t,t'} d(t,t',p)\, v^{t,t'} + \sum_{t,t'} d(t,t',b)\, v^{t,t'}$$

$$= v_N(d) + \frac{1}{2} d(s) \log N.$$

Lemma 8. *For each d,*

$$\frac{1}{|\mathcal{M}(d)|} \sum_{\mu \in \mathcal{M}(d)} \left(\prod_{i \in B(\mu)} x_i \right) \leq \prod_t x_t^{\beta_t N}.$$

Proof. Let $\mathcal{B}_t$ be a collection of subsets $B \subseteq \{i : t_i = t\}$ of type t individuals such that $|B| = \beta_t$. Symmetry implies that

$$\frac{1}{|\mathcal{M}(d)|} \sum_{\mu \in \mathcal{M}(d)} \left(\prod_{i \in B(\mu)} x_i \right) = \prod_t \frac{1}{|\mathcal{B}_t|} \sum_{B \in \mathcal{B}_t} \left(\prod_{i \in B} x_i \right) \leq \prod_t \left(\frac{1}{\alpha_t N} \sum_{i:t_i=t} x_i \right)^{\beta_t(d)N}.$$

The last inequality follows from the Newton multinomial formula. $\qquad\qquad\square$

Lemma 9. *There exist a constant $K < \infty$, such that for all N, for all $x \in [0,1]^N$*

$$\log \left[\max \left(\prod_i (1 - V_{t_i} x_i)^{J_N - 1} \prod_{\{i,j\} \in S_2 \setminus S(\mu)} \left(1 - V^{t_i,t_j} x_i x_j \right), 0 \right) \right]$$

$$\leq -J_N \sum_t V_t \alpha_t x_t - \frac{1}{2} N^2 \sum_{(t,t') \in R^*} V^{t,t'} \alpha_t x_t \alpha_{t'} x_{t'} + KN \sum_t \alpha_t x_t.$$

Proof. First, notice that $\log(1-x) \leq -x$. Thus,

$$\log \left[\prod_i (1 - V_{t_i} x_i)^{J_N - 1} \right] \leq -(J_N - 1) \sum_i V_{t_i} x_i \leq -J_N \sum_t V_t \alpha_t x_t + V^* \alpha_t x_t,$$

and

$$\log \left[\prod_{\{i,j\} \in S_2 \setminus S(\mu)} \left(1 - V^{t_i,t_j} x_i x_j \right) \right] \leq - \sum_{\{i,j\} \in S_2 \setminus S(\mu)} V^{t_i,t_j} x_i x_j$$

$$\leq - \sum_{\{i,j\} \in S_2} V^{t_i,t_j} x_i x_j + \sum_{\{i,j\} \in S(\mu) \cap S_2} V^{t_i,t_j} x_i x_j.$$

Second, observe that

$$- \sum_{\{i,j\} \in S_2} V^{t_i,t_j} x_i x_j = -\frac{1}{2} \sum_{(t,t') \in R^*} V^{t,t'} (N\alpha_t x_t)(N\alpha_{t'} x_{t'}) + \frac{1}{2} \sum_t V^{t,t} \left(\sum_{i:t_i=t} x_i^2 \right).$$

Third, because $x_i \in [0,1]$ for all i, which implies that $x_i x_{\mu(i)} \leq x_i$, we have

$$\sum_{\{i,j\} \in S(\mu) \cap S_2} V^{t_i,t_j} x_i x_j \leq V^* \sum_{\{i,j\} \in S(\mu) \cap S_2} x_i x_j \leq V^* \sum_i x_i x_{\mu(i)}$$

$$\leq V^* \sum_i x_i = V^* \sum_t N\alpha_t x_t.$$

Similarly, we show that

$$\frac{1}{2}\sum_t V_t\left(\sum_{i:t_i=t} x_i^2\right) \le \frac{1}{2}V^*\sum_t\left(\sum_{i:t_i=t} x_i\right)$$

$$\le \frac{1}{2}V^*N\sum_t \alpha_t x_t.$$

The Lemma follows from the above inequalities. $\qquad\square$

Lemma 10. *There exists a continuous increasing function* $b:[0,1]\to R_+$ *such that* $b(0)=0$ *and for each* N*, each* t

$$\prod_i (1+a(x_i)) \le \exp\left(Nb\left(\sum_t \alpha_t x_t\right)\right).$$

Proof. We assume w.l.o.g. that $a(1)>1$ (if not multiply function $a()$ by $\frac{2}{a(1)}$). For each $y>0$, let $b(y)$ be defined as a unique solution $b\ge 2a(1)y$ to the equation

$$a(1)\frac{1}{a^{-1}\left(\frac{b}{2}\right)}x = \frac{b}{2}.$$

Because the left-hand side is strictly decreasing in b and the right-hand side is strictly increasing, it is clear that the solution exists. Moreover, one shows that the solution is continuous in y, and that $\lim_{y\to 0} b(0)=0$.

Fix $x_1,...,x_N\in[0,1]$, and let $x=\frac{1}{N}\sum_i x_i$. Observe that for each c, $c\,|\{i:x_i\ge c\}|\le \sum x_i$, which implies that

$$\frac{1}{N}a^{-1}\left(\frac{b(x)}{2}\right)\left|\left\{i:x_i\ge a^{-1}\left(\frac{b(x)}{2}\right)\right\}\right| \le \frac{1}{N}\sum_i x_i = x.$$

It follows that

$$\frac{1}{N}\sum_i a(x_i) \le \frac{b(x)}{2} + \frac{1}{N}a(1)\left|\left\{i:x_i\ge a^{-1}\left(\frac{b(x)}{2}\right)\right\}\right|$$

$$\le \frac{b(x)}{2} + a(1)\frac{1}{a^{-1}\left(\frac{b(x)}{2}\right)}x = b(x).$$

The result follows from the fact that $\prod_i (1+a(x_i)) \le \exp\frac{1}{N}\sum_i a(x_i)$. $\qquad\square$

Finally, let

$$A(x) = Kx + b(x).$$

Function $A(.)$ is continuous and $A(0)=0$ by Lemma 10.

B.5. **Proof of Lemma 6: Large deviation bound.** The result follows from equation (31) and the following result.

Lemma 11. *For any continuous function $h : [0, 1]^k \to R$ of $k < \infty$ variables,*

$$\log \int_{0 \leq x_1, ..., x_k \leq 1} h(x_1, ..., x_k) \, d\mu^{m_1}(x_1) ... d\mu^{m_k}(x_k)$$

$$\leq \sum_k \log 2m_k + \sum_k m_k + \max_{0 \leq x_1, ..., x_k \leq 1} \left(\log h(x_1, ..., x_k) + \sum_k (m_k - 1) \log x_k \right).$$

Lemma 11 is a version of the Varadhan's Lemma from large deviations theory (see for example, Theorem II.7.1. in Ellis (2005)). The Varadhan's Lemma establishes asymptotic upper and lower bound on the values of certain sequences of integrals. Here, we need a proper (and not just asymptotic) upper bound.

Proof. Assume first that $k = 1$. We are going to show that

$$\int_{0 \leq x \leq 1} h(x) \, d\mu^m(x) \leq 2me^m \max_{0 \leq x \leq 1} \left(h(x) x^{m-1} \right).$$

For each probability measure μ on $[0, 1]$, any function $f : [0, 1] \to R$, let $E_\mu f = \int_0^1 f(x) \, d\mu(x)$ denote the expected value of f with respect to measure μ. Notice that

$$\mu^m(x \leq x_0) = \mu^m \left(e^{-mx/x_0} \geq e^{-m} \right)$$

$$\leq e^m E_{\mu^m} e^{-mx/x_0} \tag{35}$$

$$= e^m \left(E_\lambda e^{-x/x_0} \right)^m \leq e^m x_0^m, \tag{36}$$

where λ is the Lebesgue measure on the interval $[0, 1]$. (The first inequality is a version of Markov's inequality. The second equality comes from the fact that μ^m is a probability distribution of the average $x = \frac{1}{m} \sum x_i$ of m i.i.d. variables. The last equality comes from the fact that $E_\lambda e^{-x/x_0} = \int_0^1 e^{-\frac{1}{x_0} x} dx \leq x_0$.) Take any function $h()$ and let $h^*(x) = \max_{y \geq x} h(y)$ be the smallest decreasing function, not smaller

than $h\left(.\right)$. By integration by parts,

$$E_{\mu^m} h\left(x\right) \leq E_{\mu^m} h^*\left(x\right)$$

$$= \int_0^1 h^*\left(x\right) d\mu^m\left(x\right)$$

$$= \int_0^1 \left(-\left(h^*\right)'\left(y\right)\right) \mu^m\left(x \leq y\right) dy + h^*\left(1\right)$$

$$\leq e^m \int_0^1 \left(-\left(h^*\right)'\left(x\right)\right) x^m dx + h^*\left(1\right)$$

$$= e^m m \int_0^1 h^*\left(x\right) x^{m-1} dx - \left(e^m - 1\right) h^*\left(1\right)$$

$$\leq 2 e^m m \max_x h^*\left(x\right) x^{m-1} = 2 e^m m \max_x h\left(x\right) x^{m-1}.$$

The multi-variable case $k \geq 1$ follows from induction on k and a repeated application of single-variable case $k = 1$:

$$\int_{0 \leq x_1, \ldots, x_k \leq 1} h\left(x_1, \ldots, x_k\right) d\mu^{m_1}\left(x_1\right) \ldots d\mu^{m_k}\left(x_k\right)$$

$$= \int_{0 \leq x_k \leq 1} \left(\int_{0 \leq x_1, \ldots, x_{k-1} \leq 1} h\left(x_1, \ldots, x_{k-1}, x_k\right) d\mu^{m_1}\left(x_1\right) \ldots d\mu^{m_{k-1}}\left(x_1\right) \right) d\mu^{m_k}\left(x_k\right)$$

$$\leq \prod_{l=1}^{k-1} \left(2m_l\right) e^{\sum_{l=1}^{k-1} m_l} \int_{0 \leq x_k \leq 1} \left(\left(\max_{0 \leq x_1, \ldots, x_{k-1} \leq 1} \left(h\left(x_1, \ldots, x_{k-1}, x_k\right) \prod_{l=1}^{k-1} x_l^{m_l-1} \right) \right) \right) d\mu^{m_k}\left(x_k\right)$$

$$= \prod_{l=1}^{k-1} \left(2m_l\right) e^{\sum_{l=1}^{k-1} m_l} \left(2m_k\right) e^{m_k} \left(\max_{0 \leq x_k \leq 1} \left(\max_{0 \leq x_1, \ldots, x_{k-1} \leq 1} \left(h\left(x_1, \ldots, x_k\right) \prod_{l=1}^{k-1} x_l^{m_l-1} \right) x_k^{m_k-1} \right) \right)$$

$$= \prod_{l=1}^{k} \left(2m_l\right) e^{\sum_{l=1}^{k} m_l} \max_{0 \leq x_1, \ldots, x_k \leq 1} \left(h\left(x_1, \ldots, x_k\right) \prod_{l=1}^{k} x_l^{m_l-1} \right).$$

$\square$

B.6. **Proof of Lemma 7: Maximization.** First, notice that by changing variables $x_t = N^{-1/2} y_t$, the maximization problem in (32) turns into

$$\max_{0 \leq y_t \leq N^{1/2}} \left\{ \begin{array}{l} -j_N \sum_t V_t \alpha_t y - \frac{1}{2} \sum_{(t,t') \in R^*} V^{t,t'} \alpha_t y_t \alpha_{t'} y_{t'} \\ + \sum_t (\alpha_t + \beta_t) \log y_t + A \left(N^{-\frac{1}{2}} \sum_t \alpha_t y_t \right) \end{array} \right\} - \frac{1}{2} \sum (\alpha_t + \beta_t) \log N. \quad (37)$$

Note that $\sum (\alpha_t + \beta_t) = d(s) + d(p) + 2d(b)$.

Second, because A is a continuous function such that $\sup_{x \in [0,1]} A(x) < \infty$, and because $V^{t,t'} > 0$ for all t, t', notice that as $N \to \infty$, the y_t's that solve the maximization problem in (37) remain uniformly bounded as $N \to \infty$ (otherwise the value of the problem is dominated by the negative quadratic term.) Because $A(0) = 0$, as $N \to \infty$, the value of the maximization problem in (37) converges to

$$\to \max_{y_t \geq 0} \left\{ -j_N \sum_t V_t \alpha_t y_t - \frac{1}{2} \sum_{(t,t') \in R^*} V^{t,t'} \alpha_t y_t \alpha_{t'} y_{t'} + \sum_t (\alpha_t + \beta_t) \log y_t \right\}. \quad (38)$$

Third, we show that

Lemma 12. *For each $\delta \in D(\alpha, (\beta_t))$, the value of problem (38) is not larger than*

$$-1 - \sum_t (\alpha_t + \beta_t) \log \alpha_t - \left[\upsilon_N(\delta) + \mathcal{E}^0(\delta) + \delta(b) \right], \quad (39)$$

where

$$\mathcal{E}^0(d) = -\sum_t d(t,t,s) \log d(t,t,s)$$
$$- \frac{1}{2} \sum_{(t,t') \in R^*} (d(t,t',p) + 2d(t,t',b)) \log ((d(t,t',p) + 2d(t,t',b)) + 1).$$

Proof. Take any $\delta \in D(\alpha, (\beta_t))$. Because the expression in the square brackets of (39) is continuous in δ, it is enough to prove the lemma assuming that $\delta(t,t',c) > 0$ for each t, t', c. Observe first that for each, each type t

$$\delta(t,t,s) + \sum_{t':(t,t') \in R} (\delta(t,t',p) + 2\delta(t,t',b)) \quad (40)$$
$$= \delta(t,t,s) + \sum_{t':(t,t') \in R} (\delta(t',t,p) + 2\delta(t',t,b)) = \alpha_t + \beta_t.$$

Thus,

$$
- j_N \sum_t V_t \alpha_t y_t - \frac{1}{2} \sum_{(t,t')\in R^*} V^{t,t'} \alpha_t \alpha_{t'} y_t y_{t'} + \sum_t (\alpha_t + \beta_t) \log y_t
$$

$$
= - j_N \sum_t V_t \alpha_t y_t - \frac{1}{2} \sum_{(t,t')\in R^*} V^{t,t'} \alpha_t \alpha_{t'} y_t y_{t'} + \sum_t \delta(t,t,s) \log y_t +
$$

$$
+ \sum_{(t,t')\in R^*} \left(\frac{1}{2} \delta(t,t',p) + \delta(t,t',b) \right) \log y_t + \sum_{(t,t')\in R^*} \left(\frac{1}{2} \delta(t,t',p) + \delta(t,t',b) \right) \log y_{t'}
$$

$$
= - j_N \sum_t V_t \alpha_t y_t + \sum_t \delta(t,t,s) \log y_t
$$

$$
- \frac{1}{2} \sum_{(t,t')\in R^*} V^{t,t'} \alpha_t \alpha_{t'} y_t y_{t'} + \sum_{(t,t')\in R^*} \left(\frac{1}{2} \delta(t,t',p) + \delta(t,t',b) \right) \log y_t y_{t'}
$$

$$
= \sum_t \delta(t,t,s) \left(- \frac{j_N V_t \alpha_t}{\delta(t,t,s)} y_t + \log y_t \right)
$$

$$
+ \sum_{(t,t')\in R^*} \left(\frac{1}{2} \delta(t,t',p) + \delta(t,t',b) \right) \left(- \frac{\alpha_t \alpha_{t'} V^{t,t'}}{\delta(t,t',p) + 2\delta(t,t',b)} y_t y_{t'} + \log y_t y_{t'} \right)
$$

$$
\leq \sum_t \delta(t,t,s) \left(-1 - \log \left(\frac{j_N V_t \alpha_t}{\delta(t,t,s)} \right) \right)
$$

$$
+ \sum_{(t,t')\in R^*} \left(\frac{1}{2} \delta(t,t',p) + \delta(t,t',b) \right) \left(-1 - \log \left(\frac{\alpha_t \alpha_{t'} V^{t,t'}}{\delta(t,t',p) + 2\delta(t,t',b)} \right) \right),
$$

where the last line follows from the fact that $-ax + \log x \leq -\log a - 1$ for any $a, x > 0$. By the identity (40) and the definition of functions υ_N and $\mathcal{E}_0$, the latter is equal to

$$
= - \left(\delta(s) + \frac{1}{2}\delta(p) + \delta(b) \right) - \upsilon_N(\delta)
$$

$$
+ \sum_t \delta(t,t,s) \log \delta(t,t,s) + \frac{1}{2} \sum_{(t,t')\in R^*} (\delta(t,t',p) + 2\delta(t,t',b)) \log(\delta(t,t',p) + 2\delta(t,t',b))
$$

$$
- \sum_t \delta(t,t,s) \log \alpha_t - \sum_{(t,t')\in R^*} \left(\frac{1}{2}\delta(t,t',p) + \delta(t,t',b) \right) (\log \alpha_t + \log \alpha_{t'}).
$$

The expression in the bracket in the first line is equal to $-1 + \frac{1}{2}\delta(p)$ due to (6a). The last line is equal to $- \sum_t (\alpha_t + \beta_t) \log \alpha_t$ due to (6d), (6e) and (6a). The result follows.   □

Fourth, we show that

Lemma 13. *There exists a continuous function* $g : R \to R_+$ *such that* $g(x) > 0$ *for* $x > 0$ *and for each* $\delta \in D(\alpha, (\beta_t))$,

$$-\mathcal{E}^0(\delta) \leq -Entropy(\delta) + \delta(b) + \sum_t \beta_t \log \alpha_t - \sum_t \beta_t \log \beta_t - g(\delta(b)).$$

Proof. For each $\chi \geq 0$, let

$$l(\chi) = -\chi \log \chi.$$

Then, $l(.)$ is a concave function such that $l(1) = 0$. For future use, notice that for each $a, \chi \geq 0$,

$$l(a\chi) = \chi l(a) + a l(\chi). \tag{41}$$

Also, define

$$f(\xi) = \frac{1}{2}l(1-\xi) - \frac{1}{2}l(1+\xi) - \xi \tag{42}$$

Observe that $f(0) = 0$ and $f(.)$ is strictly decreasing and strictly concave.

For each t, t', define

$$a(t, t') = \delta(t, t', p) + \delta(t, t', b),$$

$$\xi(t, t') = \frac{1}{a(t, t')}\delta(t, t', b).$$

Then, for each $\delta \in D(\alpha, (\beta_t))$,

$$
\begin{aligned}
Entropy(\delta) &= -\frac{1}{2}\delta(p) - \delta(b) - \sum_t \delta(t, t, s) \log \delta(t, t, s) \\
&\quad + \frac{1}{2}\sum_{(t,t')\in R^*} (l(\delta(t, t', p)) + 2l(\delta(t, t', b))) \\
&= -\frac{1}{2}\delta(p) - \delta(b) - \sum_t \delta(t, t, s) \log \delta(t, t, s) \\
&\quad + \sum_{(t,t')\in R^*} \left(\frac{1}{2}l(a(t, t')(1 - \xi(t, t'))) + l(a(t, t')\xi(t, t'))\right),
\end{aligned}
$$

and

$$\mathcal{E}^0(\delta) = -\frac{1}{2}\delta(p) - \delta(b) - \sum_t \delta(t,t,s)\log\delta(t,t,s)$$

$$+ \frac{1}{2}\sum_{(t,t')\in R^*}(l(\delta(t,t',p) + 2\delta(t,t',b)))$$

$$= -\frac{1}{2}\delta(p) - \delta(b) - \sum_t \delta(t,t,s)\log\delta(t,t,s)$$

$$+ \sum_{(t,t')\in R^*}\frac{1}{2}l(a(t,t')(1 + \xi(t,t'))).$$

Moreover,

$$\sum_t \beta_t = \sum_{(t,t')\in R^*} a(t,t')\xi(t,t').$$

It follows that

$$\text{Entropy}(\delta) - \mathcal{E}^0(\delta) - \sum_t \beta_t = \sum_{(t,t')\in R^*} a(t,t')f(\xi(t,t')) + \sum_{(t,t')\in R^*} a(t,t')l(\xi(t,t')).$$

Using (41), (42), the concavity of f and l, and the fact that $\sum_{t':(t,t')\in R^*}\frac{a(t,t')}{\alpha_t}\xi(t,t') = \frac{\beta_t}{\alpha_t}$, we obtain that the latter is not larger than

$$\leq (\delta(p) + \delta(b))\sum_{(t,t')\in R^*}\frac{a(t,t')}{\delta(p) + \delta(b)}f(\xi(t,t')) + \sum_{(t,t')\in R^*} a(t,t')l(\xi(t,t'))$$

$$\leq (\delta(p) + \delta(b))f(\delta(b)) + \sum_t \alpha_t\sum_{t':(t,t')\in R^*}\frac{a(t,t')}{\alpha_t}l(\xi(t,t'))$$

$$\leq (\delta(p) + \delta(b))f(\delta(b)) + \sum_t \alpha_t l\left(\frac{\beta_t}{\alpha_t}\right)$$

$$= (\delta(p) + \delta(b))f(\delta(b)) - \sum_t \alpha_t\frac{\beta}{\alpha_t}\log\left(\frac{\beta_t}{\alpha_t}\right)$$

$$\leq \delta(b)f(\delta(b)) + \sum \beta_t\log\alpha_t - \sum_t \beta_t\log\beta_t.$$

The last equality follows from the fact that $f(\delta(b)) \leq 0$. Let $g(x) = -xf(x)$.   □

References

Ashlagi, I., J. Leshno, and Y. Kanoria (2016): "Unbalanced random matching markets: the stark effect of competition," *Journal of Political Economy*, (forthcoming). 13

Che, Y.-K., J. Kim, and F. Kojima (2015): "Stable Matching in Large Economies," . 8

Chiappori, P.-A., A. Galichon, and B. Salanie (2014): "The Roommate Problem - Is More Stable Than You Think," *Games and Economic Behavior*. 6, 18

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114(1), 175–201. 1, 3

Dagsvik, J. K. (2000): "Aggregation in Matching Markets," *International Economic Review*, 41(1), 27–58. (document), 1, 1, 2, 3, 1, 5.1, 5.1, 8

Ellis, R. (2005): *Entropy, Large Deviations, and Statistical Mechanics*. Springer, Berlin ; New York, reprint of the 1st ed. springer-verlag new york 1985 edition edn. 1, B.2, B.5

Gale, D., and L. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69(1), 9–15. 1, 2, 4

Galichon, A., and B. Salanie (2012): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," SSRN Scholarly Paper ID 1804623, Social Science Research Network, Rochester, NY. 1

Lee, S., and L. Yariv (2014): "On the Efficiency of Stable Matchings in Large Markets," SSRN Scholarly Paper ID 2464401, Social Science Research Network, Rochester, NY. 5, 15

Menzel, K. (2015): "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*. (document), 1, 4, 12, 5.1, 5.1, 8

Peski, M. (2015): "Welfare plus Entropy in Many-to-many Matching Models," Discussion paper, University of Toronto. 8

Roth, A. E., T. Sonmez, and M. U. Unver (2007): "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," *American Economic Review*, 97(3), 828–851. 1

SHAPLEY, L., AND M. SHUBIK (1971): "The assignment game I: The core," *International Journal of Game Theory*, 1(1), 111–130. 18

TAN, J. J. M. (1991): "A necessary and sufficient condition for the existence of a complete stable matching," *Journal of Algorithms*, 12(1), 154–178. 1, 2, 5, 1, 8