LARGE MATCHING MARKETS AS TWO-SIDED DEMAND SYSTEMS

KONRAD MENZEL[†]

Abstract. This paper derives asymptotic approximations to two-sided matching markets with non-transferable utility, assuming that the number of market participants grows large. We consider a model in which each agent has a random preference ordering over individual potential matching partners, and agents' types are only partially observed by the econometrician. We show that in a large market, the inclusive value is a sufficient statistic for an agent's endogenous choice set with respect to the probability of being matched to a spouse of a given observable type. Furthermore, while the number of pairwise stable matchings for a typical realization of random utilities grows at a fast rate as the number of market participants increases, the inclusive values resulting from any stable matching converge to a unique deterministic limit. We can therefore characterize the limiting distribution of the matching market as the unique solution to a fixed point condition on the inclusive values. Finally we analyze identification and estimation of payoff parameters from the asymptotic distribution of observable characteristics at the level of pairs resulting from a stable matching.

JEL Classification: C12, C13, C31, C35, C78
Keywords: Two-Sided Matching Markets, Large Games, Stable Matching, Independence of Irrelevant Alternatives

We consider identification and estimation of preference parameters in two-sided matching markets, where the researcher does not observe agents' preference rankings over matching partners, but only attributes or characteristics of the individuals involved in a transaction. Our setup assumes that the market outcome is a pairwise stable matching with non-transferable utilities (NTU), where each agent has a strict preference ordering over individuals on the opposite side of the market, and agents' types are only partially observed.

We describe our results using the language of the classical stable marriage problem. Other examples for markets without transfers between agents - or transfers that are set exogenously - include assignment of students to schools or colleges, of interns to hospitals, or contracts between workers and employers if wages or salaries are determined by centralized bargaining or the government. Markets of this type were first analyzed by Gale and Shapley (1962), and existence and properties of pairwise stable matchings are now well understood from a

Date: October 2013.
[†] NYU, Department of Economics, Email: konrad.menzel@nyu.edu
This paper has benefitted from seminar audiences at Berkeley, UCLA, Cornell, UCL, and Sciences-Po. I would especially like to thank Alexandre Belloni, Richard Blundell, Denis Chetverikov, Alfred Galichon, Ariel Pakes, and Aloysius Siow for useful conversations. Any comments or suggestions are welcome.

theoretical perspective (see Roth and Sotomayor (1990) for a summary of results). However, estimation of preference parameters from these models remains challenging due to the large number of distinct stable matchings and the interdependence between individuals' preferences and matching opportunities.

We propose a highly tractable asymptotic approximation to the distribution of matched observable characteristics resulting from pairwise stable matchings that assumes that the number of agents in the market is large. We obtain a unique limit for the type-specific match frequencies that depends on agents' preferences only through a joint surplus measure at the match level and the inclusive value for the set of matching partners available to the agent. We show that the joint surplus as a function of observable characteristics is nonparametrically identified from the asymptotic distribution, however without further assumptions, it is not possible to identify preferences on the male and female side of the market separately. For estimation of structural parameters from the limiting distribution using likelihood methods, the (endogenously determined) inclusive value functions can be treated as auxiliary parameters which solve the theoretical equilibrium conditions. We also extend the main model to the case in which individuals may only be aware of a random subset of potential matching partners.

Understanding the structure of matching markets without transfers is in itself of economic interest, and it is instructive to compare our findings to known properties of competitive markets with prices and matching markets with transfers. We find that, paralleling our intuitions for competitive markets, pairwise stability generates an essentially unique observable market outcome, where relative transaction (matching) frequencies are a function of a surplus measure at the level of a matched pair. However, in the absence of side payments or prices, the matching market clears through the relative abundance or scarcity of available matching partners of a certain type rather than explicit transfers. Furthermore, the pseudo-surplus measure captures only distributional substitution patterns between characteristics of matching partners rather than trade-offs at the individual level, and is not maximized by the market outcome. In particular, the scale of the male and female contributions to pseudo-surplus reflect the importance of systematic utility relative to idiosyncratic taste shifters in either side's random preferences rather than a common numéraire.

We characterize the stable matching as a result of every agent choosing from the set of potential spouses that are available to her or him. In particular, an available potential spouse must prefer that agent to his/her current match. Hence, the pairwise stability conditions correspond to a discrete choice problem with latent choice sets that are not observed by the researcher and endogenous outcomes of the model. Without further structure, this leads to enormous difficulties for identification analysis or estimation. In our analysis, a key simplification results from the observation that many commonly used random utility models exhibit independence of irrelevant alternatives (IIA) as a limiting property when the set of choice alternatives is large. For the conditional logit model it is known that as a result of the IIA property, an individual's matching opportunities can be summarized by the inclusive value, a scalar sufficient statistic with respect to the conditional choice probabilities.

In principle, there are many possible ways of embedding an n-agent economy into a sequence of markets. Our main objective in this paper to construct a "plausible" approximation that reflects two crucial features of the finite-agent market: For one, we intend to approximate a distribution of matching characteristics under which a nontrivial fraction of the population remains unmatched. On the other hand, we need to strike the right balance between the magnitude of observed characteristics and idiosyncratic taste shocks so that in the limit the joint distribution of matched characteristics of men and women does not degenerate to a matching rule that is deterministic or independent of observed characteristics. We show below that imposing these two requirements simultaneously results in specific rates for payoff parameters along the limiting sequence. Specifically, the limiting distribution will have the first property only if the outside option is made more attractive at exactly the right rate as the size of the market grows. The second condition concerns the scale of the distribution of unobserved heterogeneity, where the rate of increase in the variance of idiosyncratic taste shocks depends on the tail behavior of its standardized distribution.

We then show that in the limit, the equilibrium inclusive values are a deterministic function of an individual's observable attributes alone. Furthermore, the inclusive value functions generated by a stable matching are uniquely determined by the marginal distributions of men's and women's observable characteristics. This result does not imply uniqueness of the stable matching in the limit - in fact it is known that the typical number of stable matchings increases at an exponential rate in the number of participants. Rather, the increasing number of distinct matchings are in the limit indistinguishable from the researcher's perspective since they all result in the same limiting joint distribution of matched observable characteristics.

Related Literature. Logan, Hoff, and Newton (2008) estimate a model for a matching market, and Christakis, Fowler, Imbens, and Kalyanaraman (2010) use MCMC techniques to estimate a model of strategic network formation, but do not explicitly account for the possibility of multiple equilibria. The matching market model analyzed in this paper is different from that in Choo and Siow (2006), Fox (2010) and Galichon and Salanié (2012) in that we do not assume transferable utilities, so that stable matchings do not necessarily maximize joint surplus across matched pairs. Also, we allow for unobserved heterogeneity in preferences over individual potential spouses on the other side of the market. In contrast, previous work by Choo and Siow (2006), Galichon and Salanié (2012), and Hsieh (2012) assumes that unobserved taste shifters are group-specific, i.e. agents are indifferent between potential spouses of the same observable type. Decker, Lieb, McCann, and Stephens (2013) give a representation of equilibrium in the transferable utilities with finitely many types in terms of the share of unmarried agents of each type, and show uniqueness of the resulting matching equilibrium.

Echenique, Lee, and Shum (2010) and Echenique, Lee, Shum, and Yenmez (2012) consider inference based on implications of matching stability assuming that agents' types are discrete and fully observed by the econometrician. Pakes, Porter, Ho, and Ishii (2006), Baccara, Imrohoroglu, Wilson, and Yariv (2012) and Uetake and Watanabe (2012) estimate matching games via inequality restrictions on the conditional mean or median of payoff functions derived from necessary conditions for optimal choice. Instead, we model the full distribution of payoffs and match characteristics. In general, our approach requires some conditions on the distribution of unobserved heterogeneity (most importantly independence from observed characteristics and across agents), but on the other hand that knowledge permits to compute policy-relevant counterfactuals (e.g. conditional choice probabilities) based on estimated parameters and the asymptotic approximation to the distribution of matched characteristics.

Our main approach towards identification and estimation is to work with a more tractable large-sample approximation to the distribution of observable characteristics at the level of pairs resulting from a stable matching. Our formal derivations for the limiting matching games rely heavily on insights from Dagsvik (2000)'s analysis of aggregate matchings. Specifically, Dagsvik (2000) showed convergence of the distribution of matchings with no observable attributes under the assumption that individual taste shifters are extreme value type-I distributed and private knowledge, and characterized the limiting distribution for discrete attributes, assuming convergence. In contrast to our analysis, his model imposes a market clearing condition with respect to type-specific supply and demand probabilities as a solution concept, whereas we assume pairwise stability given the realized values of non-transferable utilities. We also extend his arguments to the empirically relevant case of continuous covariates, establishing uniqueness of the limiting distribution of observable characteristics in matched pairs. Previous characterizations of the distribution of match characteristics in both the TU and the NTU setting were limited to the case of finitely many observable types (this includes the results by Dagsvik (2000), Choo and Siow (2006), Galichon and Salanié (2012), Hsieh (2012), and Graham (2013)). To our knowledge the only exception is Dupuy and Galichon (2013)'s continuous extension to Choo and Siow (2006)'s characterization of equilibrium in the TU matching model.

Agarwal (2012) and Azevedo and Leshno (2012) consider limiting distributions as the number of individuals on one side of the market grows. Interestingly, in both models, the stable matching in the limiting game is unique. However, in our analysis the number of agents on both sides of the market grows, and although we show that the limiting distribution of matched characteristics is unique, the number of distinct stable matchings grows at a fast rate. The stable matchings will in general differ from the agents' perspective even in the limit - both in terms of welfare and the identity of their match - and our uniqueness result only concerns observational equivalence from the econometrician's perspective.

The implications of the independence of irrelevant alternatives (IIA) property for discrete choice models with a finite, or countably infinite number of alternatives have been analyzed by R.Luce (1959), McFadden (1974), Yellott (1977), Cosslett (1988), Resnick and Roy (1991), and Dagsvik (1994). We show that for the NTU model, the pairwise stability conditions translate into a discrete choice problem with unobserved and endogenous choice sets containing a large number of alternatives, where IIA arises as a limiting property of conditional choice probabilities, which greatly simplifies our analysis.

Notation. We use standard "little-o"/"big-O" notation to denote orders of convergence for deterministic sequences, and convergence in probability for random sequences. We also write $a_n \asymp b_n$ if $\lim_n a_n/b_n = 1$. We also use the abbreviations a.s. and w.p.a.1 for the qualifiers "almost surely" and "with probability approaching one."

2. Model for Finite Economy

The researcher observes data on matching outcomes from one or several markets. We assume that a data set contains variables x_i and z_j that have information on some of women i's and man j's characteristics, where both vectors may contain discrete and continuous variates. We denote the number of women and men in the market with n_w and n_m, respectively. The marginal distributions of x_i and z_j in the population are given by the p.d.f.s $w(x)$ and $m(z)$, and we denote the supports of x_i and z_j with $\mathcal{X}$ and $\mathcal{Z}$, respectively.

Furthermore, we observe the outcome of the matching process, where we have information whether man j and woman i are matched, and which individuals remain single. Specifically, we use $\mu_w(i)$ and $\mu_m(j)$ to denote woman i's, and man j's spouse, respectively, under the matching μ. Each individual can marry a person of the opposite sex or choose to remain single.

2.1. Types and Preferences.

We consider a matching model with non-transferable utilities (NTU), where preferences over spouses are given by the latent random utility functions of the form

$$\begin{aligned} U_{ij} &= U(x_i, z_j) + \sigma\eta_{ij} \\ V_{ji} &= V(z_j, x_i) + \sigma\zeta_{ji} \end{aligned} \tag{2.1}$$

for $i = 1, \ldots, n_W$ and $j = 1, \ldots, n_M$. The random utility for the outside option - i.e. of remaining single - is specified as

$$\begin{aligned} U_{i0} &= 0 + \sigma \max_{k=1,\ldots,J} \{\eta_{i0,k}\} \\ V_{j0} &= 0 + \sigma \max_{k=1,\ldots,J} \{\zeta_{j0,k}\} \end{aligned} \tag{2.2}$$

where for our asymptotic sequence, J is assumed to increase at a rate to be specified below, and the scale parameter σ may also depend on market size.

In our setup, payoffs have "systematic" components $U(x_i, z_j)$ and $V(z_j, x_i)$ that are a function of individual i and j's characteristics x_i, and z_j, respectively, and the "idiosyncratic" components η_{ij}, ζ_{ji} and $\eta_{i0,k}, \zeta_{j0,k}$ are i.i.d. draws from a known distribution that are independent of $x_1, x_2, \dots$ and $z_1, z_2, \dots$. For our derivation of the limiting distribution, we do not explicitly distinguish between components of x_i and z_j that are observed or unobserved. However the central difference between the roles of the "systematic" and the "idiosyncratic" parts of the payoff functions in our model is that random taste shifters η_{ij} and ζ_{ji} are assumed to be independent across i, j and therefore do not induce correlation in preferences across agents. The appropriate choice of a sequence for σ will generally depend on the shape of the tails of $G(\cdot)$ and will be discussed below.

The model differs from the assumptions in previous work by Choo and Siow (2006) and Galichon and Salanié (2012) in that the idiosyncratic taste shocks for woman i and man j, η_{ij} and ζ_{ji}, respectively, are individual-specific with respect to potential spouses $j = 1, \dots, n_m$ and $i = 1, \dots, n_w$ rather than only allowing for heterogeneity in tastes over a finite number of observable characteristics. In that aspect, our random utility model is similar to that in Dagsvik (2000) - however our setup also allows for the systematic part of the random utility functions to depend on continuous characteristics, and we do not assume a particular distribution for the idiosyncratic taste shifters.

The rationale for modeling the outside option as the maximum of J independent draws for the idiosyncratic taste shifters is that as the market grows, the typical agent can choose from an increasing number of potential spouses. Since in our setup the shocks η_{ij} and ζ_{ji} generally have unbounded support, any alternative with a fixed utility level will eventually be dominated by one of the largest draws for the increasing set of potential matching partners. Hence, by allowing the agent to sample an increasing number of independent draws for the outside option, it can be kept sufficiently attractive to ensure that the share of unmatched agents remains stable along the sequence. Alternatively, one could model the outside option as $\tilde{U}_{i0} = \log J + \sigma \eta_{i0}$, where η_{i0} is a single draw from the distribution $G(\eta)$, as e.g. in Dagsvik (2000). Preliminary calculations suggest that both approaches lead to equivalent results, but the formulation with "multiple outside options" is more convenient for our derivations than the "location shift" version.

The assumption of non-transferable utility makes our results applicable to markets in which transfers between matching partners are restricted or ruled out altogether. Institutional restrictions on transfers or side payments are often motivated by ethical or distributional concerns, and are common e.g. for assigning students to schools or colleges or residents to hospitals in the medical match. Note that our assumptions allow for transfers that are deterministic functions of characteristics x_i, z_j, in which case we can interpret $U(x_i, z_j)$ and $V(z_j, x_i)$ as the systematic parts of payoffs net of transfers. For example, employment contracts between workers and a firms may be subject to collective bargaining agreements, which stipulate a fixed wage given the job description and worker's education, experience, or tenure at the firm.

Throughout the paper, we will maintain that the deterministic parts of random payoffs satisfy certain uniform bounds and smoothness restrictions:

Assumption 2.1. ***(Systematic Part of Payoffs)*** *The functions $|U(x,z)| \leq \bar{U} < \infty$ and $|V(z,x)| \leq \bar{V} < \infty$ are uniformly bounded in absolute value and continuous in $\mathcal{X} \times \mathcal{Z}$. Furthermore, at all $(x', z')' \in \mathcal{X} \times \mathcal{Z}$ the functions $U(x,z)$, and $V(z,x)$ are $p \geq 0$ times differentiable with uniformly bounded partial derivatives.*

For notational simplicity, in the formulation of the differentiability requirements we do not distinguish between components of x, z that are continuous and those which only take discrete values. The condition could be weakened to hold only with respect to the continuous parts, holding the discrete components fixed.

We next state our assumptions on the distribution of unobserved taste shifters. Most importantly, we impose sufficient conditions for the distribution of $\max_j \eta_{ij}$ to belong to the domain of attraction of the extreme-value type I (Gumbel) distribution. Following Resnick (1987), we say that the upper tail of the distribution $G(\eta)$ is of type I if there exists an auxiliary function $a(s) \geq 0$ such that the c.d.f. satisfies

$$\lim_{s\to\infty} \frac{1 - G(s + a(s)v)}{1 - G(s)} = e^{-v}$$

for all $v \in \mathbb{R}$. We are furthermore going to restrict our attention to distributions for which the auxiliary function can be chosen as $a(s) := \frac{1-G(s)}{g(s)}$. We can now state our main assumption on the distribution of the idiosyncratic part of payoffs:

Assumption 2.2. ***(Idiosyncratic Part of Payoffs)*** *η_{ij} are i.i.d. draws from the distributions $G(s)$, and are independent of x_i, z_j, where (i) the c.d.f. $G(s)$ is absolutely continuous with density $g(s)$, and (ii) the upper tail of the distribution $G(s)$ is of type I with auxiliary function $a(s) := \frac{1-G(s)}{g(s)}$.*

As we already pointed out before, conditional independence of idiosyncratic taste shifters across alternatives is a strong, but nevertheless very common restriction in the discrete choice literature. Conditional independence implies that any correlations in preferences for different alternatives must be due to the systematic part of random utility functions. In order to illustrate the main difference between our NTU setup and the alternative assumption of transferable utilities, note that transfers resulting from splitting the match surplus mechanically depend on the idiosyncratic taste shifters and therefore generally result in correlated net payoffs U_{ij} and V_{ji} conditional on x_i, z_j. Hence, while we can easily extend our framework to allow for transfers that are deterministic functions of x_i, z_j, the assumption of conditional independence of net payoffs U_{ij} and V_{ji} cannot be reconciled with the standard transferable utility model.

Continuity of the distribution of η in part (i) is fairly standard, and as discussed above, parts (ii) of the assumption ensure that the distribution of unobserved taste shocks belongs to the domain of attraction of the extreme-value type-I (Gumbel) distribution. While we do not give more primitive conditions for part (iii), it is possible to show that it is satisfied for most parametric specifications for the distribution of η_{ij} commonly used in discrete choice models. E.g. for the Extreme-Value Type-I distribution, part (a) holds with auxiliary function $a(s) = 1$, whereas for standard normal taste shifters, the condition holds with auxiliary function $a(s) = \frac{1}{s}$.[1]

2.2. **Pairwise Stability.** One standard solution concept for this market is matching stability: A matching μ is *pairwise stable* if given the preferences $(U_{ij}, V_{ji})_{ij}$, every individual prefers her/his spouse under μ to any other achievable partner. I.e. μ must satisfy the conditions (i) if $U_{ij} > U_{i\mu_w(i)}$, then $V_{j\mu_m(j)} > V_{ji}$, and (ii) if $V_{ji} > V_{j\mu_m(j)}$, then $U_{i\mu_w(i)} > U_{ij}$.

The problem of stable matchings with non-transferable utility has been studied extensively and is well-understood from a theoretical perspective.[2] If preferences are strict, a stable matching always exists, and in general the set of stable matchings has a minimal element μ_M and a maximal element μ_W with respect to the preferences of the female side which we refer to as the W-preferred and M-preferred stable matching, respectively, whereas the preferences over matchings on the male side are exactly opposed. Specifically, for any stable matching μ^*, the W-preferred stable matching satisfies $U_{i\mu_W(i)} \geq U_{i\mu^*(i)}$ and $V_{j\mu_W(j)} \leq V_{j\mu^*(j)}$, and for the M-preferred stable matching we always have $U_{i\mu_M(i)} \leq U_{i\mu^*(i)}$ and $V_{j\mu_M(j)} \geq V_{j\mu^*(j)}$. The M-preferred and W-preferred stable matching can be found in polynomial computing time using the well-known Gale-Shapley algorithm. The number of distinct stable matchings for a typical realization of preferences increases exponentially in the number of individuals on each side of the market (see Theorem 3.19 in Roth and Sotomayor (1990)).

Pairwise stable matchings may arise in centralized matching markets that employ a variant of the Gale-Shapley algorithm to assign matching partners. Roth and Vande Vate (1990) show that randomized myopic tâtonnement processes converge to pairwise stable matchings with probability one if allowed to continue indefinitely, which may also justify matching stability as a suitable empirical model for decentralized markets. It is also important to notice that pairwise stability as a solution concept does not require that agents have perfect knowledge of all participants' preferences, but each agent only needs to know which matching

[1] See Resnick (1987), pages 42-43.

[2] See Roth and Sotomayor (1990) for a synthesis of the classical results on the problem.

partners are available to him or her. In section 6, we also give a straightforward extension to the main model which allows for agents only to be aware of a random subset of potential matching partners.

For our analysis it is useful to translate the pairwise stability conditions into a discrete-choice problem at the individual level: Given a matching μ, we let the set $M_i \equiv M_i[\mu] \subset \{0, 1, \ldots, n_M\}$ denote the set of men j preferring woman i over their current match, $\mu(j)$, i.e.

$$j \in M_i[\mu] \text{ if and only if } V_{ji} \geq V_{j\mu(j)}$$

We call $M_i[\mu]$ the set of men available to woman i under the matching μ, or woman i's *opportunity set*, and by default we assume that $0 \in M_i$, i.e. the opportunity set always includes the outside option. Similarly, we define man j's opportunity set $W_j \equiv W_j[\mu]$ as

$$i \in W_j[\mu] \text{ if and only if } U_{ij} \geq U_{i\mu(i)}$$

where we assume $0 \in W_j$. For the following it is useful to define

$$U_i^*(M) := \max_{j \in M} U_{ij} \text{ and } V_j^*(W) := \max_{i \in W} V_{ji}$$

Since $0 \in M_i$, we have in particular that $U_i^*(M_i) \geq 0$. It is straightforward to verify that the matching μ is stable iff

$$U_{i\mu(i)} \geq U_i^*(M_i) \text{ and } V_{j\mu(j)} \geq V_j^*(W_j) \tag{2.3}$$

for all $i = 1, \ldots, n_w$ and $j = 1, \ldots, n_m$.

2.3. **Inclusive Values.** The main difficulty in solving for the distribution of stable matchings consists in the fact that the set of available spouses M_i and W_j for each individual is not observed by the researcher, and determined endogenously in the market. However, the problem simplifies substantially if idiosyncratic taste shifters are independent of the opportunity sets, and extreme-value type-I distributed so that the resulting conditional choice probabilities satisfy the independence of irrelevant alternatives (IIA) property. In that the composition and size of the set of alternatives affects the conditional choice probabilities only through the inclusive value, a scalar parameter summarizing the systematic components of payoffs for the available options, see R.Luce (1959), McFadden (1974), and Dagsvik (1994). We show below that under Assumption 2.2, the conditional choice probabilities converge to those implied by the conditional logit (extreme-value type I) model in the limit. Most importantly we find that in that limit, the respective inclusive values of the (endogenous) opportunity sets M_i and W_j are sufficient statistics with respect to the conditional choice probabilities, and pairwise stability is asymptotically equivalent to an equilibrium condition on the inclusive values.

Holding a woman i's opportunity set fixed exogenously at M, define the conditional probability that woman i is available to man j given characteristics x_i, z_j as

$$\Lambda_w(x, z; M) := P(U_{ij} \geq U_i^*(M) | z_j = z, x_i = x) \tag{2.4}$$

Similarly, for a fixed opportunity set W for man j, the conditional probability that he is available to woman i given characteristics x_i, z_j is given by

$$\Lambda_m(z, x; W) := P(V_{ji} \geq V_j^*(W) | x_i = x, z_j = z)$$

Note that $\Lambda_m(\cdot; W)$ and $\Lambda_w(\cdot; M)$ do in general not coincide with the conditional probabilities given the equilibrium values of W_j and M_i, respectively, but are structural choice probabilities given an exogenously determined opportunity set.[3]

We also define the *inclusive values* for woman i's opportunity set M, $I_w[M]$, and for man j's opportunity set W, $I_m[W]$ respectively, as

$$\begin{aligned} I_{wi}[M] &:= \frac{1}{J} \sum_{j \in M} \exp\{U(x_i, z_j)\} \\ I_{mj}[W] &:= \frac{1}{J} \sum_{i \in W} \exp\{V(z_j, x_i)\} \end{aligned}$$

The normalization by J^{-1} is arbitrary at this point, but will be convenient for the asymptotic analysis. Specifically, we show below that under our assumptions, the size of a typical participant's opportunity set grows at a rate proportional to J, and the inclusive value converges to a deterministic limit. Note also that we depart from the usual definition of the inclusive value as the conditional expectation of woman i's indirect utility from a choice set M,

$$\begin{aligned} \mathbb{E}\left[\max_{j \in M} U_{ij} \middle| x_i, (z_j)_{j \in M}\right] &= \log\left(J + \sum_{j \in M} \exp\{U(x_i, z_j)\}\right) + \kappa \\ &= \log(I_{wi}[M] - 1) + \log J + \kappa \end{aligned}$$

where κ is Euler's constant, see e.g. McFadden (1974). While there is a one-to-one transformation between the inclusive value according to the usual definition and ours, we find it more convenient to describe our results in terms of the variables $I_{wi}[M]$ and $I_{mj}[w]$.

[3]Here, we use the term "structural" in the sense of the structural form in a simultaneous equations model, see e.g. Hausman (1983), as opposed to conditional choice probabilities given endogenously determined sets M_i and W_j, which would correspond to a "reduced form" of the model.

Now, given our setup and assuming IIA,

$$\begin{aligned}\Lambda_m(z_j, x_i; W) &= \frac{\exp\{V(z_j, x_i)\}}{J + \sum_{i' \in W \cup \{i\}} \exp\{V(z_j, x_{i'})\}} \\ &= \frac{J^{-1} \exp\{V(z_j, x_i)\}}{1 + I_{mj} + J^{-1} \mathbb{1}\{i \notin W\} \exp\{V(z_i, x_i)\}}\end{aligned}$$

Now note that $I_{mj} \geq 0$ almost surely, so that the denominator of the expression on the right-hand side is bounded away from zero. Since $|V(z_j, x_i)| \leq \bar{V}$ is bounded, we can approximate $\Lambda_m(z_j, x; W)$ by

$$\Lambda_m(z_j, x; I_{mj}) := \frac{J^{-1} \exp\{V(z_j, x)\}}{1 + I_{mj}} = \Lambda_m(z_j, x; W) + o_p\left(J^{-1}\right)$$

as J becomes large, where the second equality follows from the continuous mapping theorem. We can now denote the expectation of woman i's inclusive value after setting men's inclusive values to I_{mj}, $j = 1, \dots, n_m$, exogenously by

$$\begin{aligned}\hat{\Gamma}_{wi}(I_{m1}, \dots, I_{mn_m}) &:= \frac{1}{J} \sum_{j=1}^{n_m} \exp\{U(x_i, z_j)\} \Lambda_m(z_j, x_i; I_{mj}) \\ &= \frac{1}{J^2} \sum_{j=1}^{n_m} \frac{\exp\{U(x_i, z_j) + V(z_j, x_i)\}}{1 + I_{mj}}\end{aligned}$$

Note that the right-hand side expression is symmetric with respect to permutations of the tuples (z_j, I_{mj}) and depends only on woman i's attributes x_i. Hence we define

$$\begin{aligned}\hat{\Gamma}_w(x) &:= \frac{1}{J^2} \sum_{j=1}^{n_m} \frac{\exp\{U(x, z_j) + V(z_j, x)\}}{1 + I_{mj}} \\ \hat{\Gamma}_m(z) &:= \frac{1}{J^2} \sum_{i=1}^{n_w} \frac{\exp\{U(x_i, z) + V(z, x_i)\}}{1 + I_{wi}}\end{aligned}$$

For the discussion below, it is also important to notice that the functions $\hat{\Gamma}_w(x)$ and $\hat{\Gamma}_m(z)$ are defined for all values of $x \in \mathcal{X}$ and $z \in \mathcal{Z}$, respectively, and not only the values of x_i and z_j that are seen in the market. Furthermore, since $I_{mj} \geq 0$ and $I_{wi} \geq 0$ a.s., $\hat{\Gamma}_w(x)$ and $\hat{\Gamma}_m(z)$ inherit boundedness, continuity and other smoothness properties from the functions $U(\cdot)$ and $V(\cdot)$.

3. Limiting Model with a Continuum of Agents

We now define a distributional game which will later be shown to be the appropriate limit for certain sequences of the finite agent version of the two-sided matching market. Although for the purposes of this paper, we are primarily interested in the limit of the matching market as an approximation to the finite agent problem, we argue that it can be motivated independently as a random utility model of spousal choice with a continuum of agents. We rely on work by Cosslett (1988), Resnick and Roy (1991) and Dagsvik (1994), for a coherent formulation of a random utility model with infinitely many alternatives, and we refer the reader to these papers for specific definitions and results.

Following Resnick and Roy (1991) and Dagsvik (1994), suppose that spousal preferences can be represented by random upper semi-continuous (USC) utility functions satisfying the independence of irrelevant alternatives (IIA) property. The USC property is imposed to ensure the existence of maxima on compact choice sets. Suppose there is a mass $\exp\{\gamma_w\}$ of women in the market with types distributed according to the p.d.f. $w(x)$, and a mass $\exp\{\gamma_m\}$ of men with types drawn from $m(z)$, where $|\gamma_w|, |\gamma_m| \leq \gamma^* < \infty$. Also assume that the matching opportunities for a woman of type $x_i = x$ are given by a random opportunity set M which is generated by a Poisson process with intensity measure that is a mixture of a density $m(z, x)$ and a unit mass for the outside option. Furthermore, we assume that M is independent of woman i's preferences. Similarly, we let $w(z, x)$ be the Poisson measure generating the opportunity set for a male of observable type $z_j = z$. For simplicity, we also assume that $w(z, x)$ and $m(x, z)$ have full support on $\mathcal{X} \times \mathcal{Z}$. Furthermore, we let $z^*(M)$ denote the vector of observable characteristics of the (random) arg max over M. Note that from results by Resnick and Roy (1991), this argmax is well-defined with probability one.

By Theorem 4 in Dagsvik (1994), the independence of irrelevant alternatives (IIA) condition together with additive separability of the random utility functions implies that

$$P\left(z^*(M) \leq z | x_i = x\right) = \frac{\int_{s\leq z} \exp\{U(x, s)\} m(s, x) ds}{1 + \int \exp\{U(x, s)\} m(s, x) ds} \equiv \frac{\int_{s\leq z} \exp\{U(x, s)\} m(s, x) ds}{1 + \Gamma_w[m](x)} \tag{3.1}$$

In analogy to the finite-sample matching market, we call

$$\Gamma_w[m](x) := \int \exp\{U(x, s)\} m(s, x) ds$$

the *inclusive* value of the opportunity set given by $m(z, x)$ to a woman of type x. Similarly, we let

$$\Gamma_m[w](z) := \int \exp\{V(z, s)\} w(s, z) ds$$

denote the *inclusive* value of the opportunity set given by $w(x, z)$ to a man of type z.

Next, we derive equilibrium conditions for the Poisson intensities $w(x, z)$ and $m(z, x)$: let $w(z|x)$ to denote the conditional probability that a woman i with observable characteristics $X_i = x$ chooses a man j with characteristics $Z_j = z$ over the best alternative in her opportunity set M. Differentiating the joint c.d.f. in (3.1) with respect to z and dividing by $m(z, x)$, we find that

$$w(z|x) = \frac{\exp\{U(x, s)\}}{1 + \Gamma_w[m](x)}$$

Note that following the same reasoning as in Section 2.3, this choice probability does not depend on whether man j is in M or not. Using the same steps for the other side of the market, and noting that $w(z,x) = w(z|x)w(x)\exp\{\gamma_w\}$, we obtain the following equilibrium conditions on choice probabilities as a generalization of the results in Theorem 3 in Dagsvik (2000):

$$\begin{aligned} \frac{w(x,z)}{w(x)\exp\{\gamma_w\}} &= \frac{\exp\{U(x,z)\}}{1+\int \exp\{U(x,s)\}m(s,x)ds} \\ \frac{m(z,x)}{m(z)\exp\{\gamma_m\}} &= \frac{\exp\{V(z,x)\}}{1+\int \exp\{V(z,s)\}w(s,z)ds} \end{aligned} \tag{3.2}$$

For our analysis it turns out to be more convenient to formulate the equilibrium conditions in terms of the inclusive value functions $\Gamma_w[m](x) := \Gamma_w(x)$ and $\Gamma_m(z) := \Gamma_m[w](z)$,

$$\begin{aligned} \Gamma_w(x) &= \int \frac{\exp\{U(x,s)+V(s,x)+\gamma_m\}m(s)}{1+\Gamma_m(s)}ds \\ \Gamma_m(z) &= \int \frac{\exp\{U(s,z)+V(z,s)+\gamma_w\}w(s)}{1+\Gamma_w(s)}ds \end{aligned} \tag{3.3}$$

where we substituted the p.d.f.s from 3.2 into the definitions of $\Gamma_w[m](\cdot)$ and $\Gamma_m[w](\cdot)$. In order to simplify notation, we define the operators $\Psi_m : \Gamma_w \mapsto \Gamma_m$ and $\Psi_w : \Gamma_m \mapsto \Gamma_w$ by

$$\begin{aligned} \Psi_w[\Gamma_m](x) &:= \int \frac{\exp\{U(x,s)+V(s,x)+\gamma_m\}m(s)}{1+\Gamma_m(s)}ds \\ \Psi_m[\Gamma_w](z) &:= \int \frac{\exp\{U(s,z)+V(z,s)+\gamma_w\}w(s)}{1+\Gamma_w(s)}ds \end{aligned} \tag{3.4}$$

In that notation, we can rewrite the equilibrium conditions in (3.3) as the fixed-point problem

$$\Gamma_w^* = \Psi_w[\Gamma_m^*] \quad \text{and } \Gamma_m^* = \Psi_m[\Gamma_w^*] \tag{3.5}$$

To simplify notation, in the following we will write $\Gamma := (\Gamma_w, \Gamma_m)$ and $\Psi[\Gamma] := (\Psi_w[\Gamma_m], \Psi_m[\Gamma_w])$. We then consider the equivalent problem

$$\log \Gamma = \log \Psi[\Gamma]$$

and show that under Assumption 2.1, the mapping $\log \Gamma \mapsto \log \Psi[\Gamma]$ is a contraction, implying existence and uniqueness of the equilibrium inclusive value function Γ^*.

Theorem 3.1. ***(Unique Equilibrium in Limiting Game)*** *Under Assumption 2.1, (i) the mapping* $(\log \Gamma) \mapsto (\log \Psi[\Gamma])$ *is a contraction mapping with*

$$\left\|\log \Psi[\Gamma] - \log \Psi[\tilde{\Gamma}]\right\| \leq \lambda \left\|\log \Gamma - \log \tilde{\Gamma}\right\|$$

where $\lambda := \frac{\exp\{\bar{U}+\bar{V}+\gamma^*\}}{1+\exp\{\bar{U}+\bar{V}+\gamma^*\}} < 1$ *and does not depend on* θ*. Specifically, a solution to the fixed point problem in (3.5) exists and is unique. (ii) Moreover the equilibrium distributions are* *characterized by functions $\Gamma_w^*(x)$ and $\Gamma_m^*(z)$ that are continuous and p times differentiable in x and z, respectively, with bounded partial derivatives.*

See the appendix for a proof. The second part of the theorem essentially says that the inclusive value functions $\Gamma_w^*(x)$ and $\Gamma_m^*(z)$ solving the limiting problem are very smooth. This is a straightforward implication of the observation that the respective ranges of the operators Ψ_w and Ψ_m are classes of functions with the same smoothness properties as the payoff functions $U(x,z)$ and $V(z,x)$.

It is important to note that uniqueness of the limiting inclusive value functions can be easily reconciled with a growing number of stable matchings as the number of agents in the market increases: If the stable matching is not unique, then by standard results from the theory of two-sided matching markets with non-transferable utilities, the number of available men in the women-preferred matching has to be weakly larger than in the men-preferred matching for all women, and strictly larger for at least one woman. We do in fact find in simulations that the number of individuals for whom the opportunity sets differ between the extremal matchings diverges to infinity as the market grows. However that difference grows more slowly than the typical size of the opportunity sets under any matching, so that under the scale normalization in the definition of $\hat{\Gamma}_w$ and $\hat{\Gamma}_m$, the gap between the inclusive values corresponding to the M-preferred and W-preferred matching vanishes as the number of individuals in the economy grows large.

We can now characterize the matching outcome in terms of the measure F of matched characteristics corresponding the number/mass of couples with observable characteristics x, z resulting from the matching. Note that this measure does not correspond to a probability distribution but integrates to a value between $\exp\{\max\{\gamma_w, \gamma_m\}\}$ and $\exp\{\gamma_w\} + \exp\{\gamma_m\}$ that may be different from one. However, we can use this measure to derive the sampling distributions of matched characteristics for any given sampling protocol, e.g. depending on whether the researcher selects individuals or couples at random. Implications for identification and estimation of preference parameters will be discussed in more detail in section 5.

Let $f(x,z)$ denote the joint density of observable characteristics of a matched pair, defined as the Radon Nikodym derivative of the measure F. As a convention we let $f(x,*)$ and $f(*,z)$ denote the density of characteristics among unmatched women and men, respectively. In particular, the following relations hold:

$$\int_{\mathcal{Z}} f(x,z)dz + f(x,*) = w(x)\exp\{\gamma_w\} \quad \text{and} \quad \int_{\mathcal{X}} f(x,z)dx + f(*,z) = m(z)\exp\{\gamma_m\}$$

Then, given the inclusive value functions $\Gamma_w^*(x)$ and $\Gamma_m^*(z)$ characterizing the equilibrium opportunity sets, we obtain the (mass) distribution

$$\begin{aligned} f(x,z) &= \frac{\exp\{U(x,z)+V(z,x)+\gamma_w+\gamma_m\}w(x)m(z)}{(1+\Gamma_w^*(x))(1+\Gamma_m^*(z))} \quad x\in\mathcal{X}, z\in\mathcal{Z} \\ f(x,*) &= \frac{w(x)\exp\{\gamma_w\}}{1+\Gamma_w^*(x)} \quad x\in\mathcal{X} \\ f(*,z) &= \frac{m(z)\exp\{\gamma_m\}}{1+\Gamma_m^*(z)} \quad z\in\mathcal{Z} \end{aligned} \tag{3.6}$$

where Γ_w^* and Γ_m^* solve the fixed-point equation in 3.5.

4. Asymptotic Approximation for Finite Market

This section establishes convergence of the distributions generated by stable matchings in the finite economy to the limiting model in (3.6) derived in the previous section. The asymptotic argument will proceed in four main steps: First, we show convergence of conditional choice probabilities (CCP) to CCPs generated by the Extreme-Value Type-I taste shifters, under the assumption that taste shifters η_{ij} are independent from the equilibrium opportunity sets W_i and M_j. Secondly, we demonstrate that dependence of taste shifters and opportunity sets is negligible for CCPs when n is large. Hence we can approximate choice probabilities using the inclusive values. In a third step, we establish that the inclusive values $I_w[M_i]$ and $I_m[W_j]$ are approximated by their conditional means $\hat{\Gamma}_w(x_i)$ and $\hat{\Gamma}_m(z_j)$. Finally we show that the inclusive value functions $\hat{\Gamma}_w$ and $\hat{\Gamma}_m$ corresponding to stable matchings are approximate solutions to a sample analog of the fixed-point problem $\Gamma = \Psi_0[\Gamma]$. Uniform convergence of the sample fixed point mapping to its population version then implies that the inclusive values converge in probability to the limits Γ_w^* and Γ_m^*.

Next, we specify the limiting sequence of markets. In general, there are many different ways in which we could embed the n-player economy into an asymptotic sequence. However, we want the approximation to retain the following features of the finite-agent market: for one, the share of single individuals should not degenerate to one or zero. Furthermore, we want the systematic parts of payoffs to remain predictive for match probabilities in the limit, although the joint distribution of male/female match characteristics should also not be degenerate in the limit. For the first requirement, it is necessary to increase the payoff from outside option as the number of available alternatives grows, whereas to balance the relative scales of the systematic and idiosyncratic parts we have to choose the scale parameter $\sigma \equiv \sigma_n$ at an appropriate rate.

Assumption 4.1. ***(Market Size)*** *(i) The size of a given market is governed by $n = 1, 2, \ldots,$ where the number of men and women $n_m = n\exp\{\gamma_m\}$ and $n_w = n\exp\{\gamma_w\}$ grow proportionally with n, where γ_w and γ_m are bounded in absolute value across markets. (ii)*

The size of the outside option is $J = \left[n^{1/2}\right]$*, where* $[x]$ *denotes the value of* x *rounded to the closest integer. (iii) The scale parameter for the taste shifters* $\sigma \equiv \sigma_n = \frac{1}{a(b_n)}$*, where* $b_n = G^{-1}\left(1 - \frac{1}{\sqrt{n}}\right)$*, and* $a(s)$ *is the auxiliary function specified in Assumption 2.2 (ii).*

Part (i) requires that the number of men and women in each market is of comparable magnitude. The rate for J in part (ii) is chosen to ensure that the share of unmatched agents is bounded away from zero and one along the sequence.[4]

The construction of the sequence σ_n in part (iii) implies a scale normalization for the deterministic parts $\tilde{U}_{ij}$, and is chosen as to balance the relative magnitude for the respective effects of observed and unobserved taste shifters on choices as n grows large. Specifically, for an alternative rate $\tilde{\sigma}_n$ such that $\tilde{\sigma}_n a(b_n) \to 0$, the systematic parts of payoffs $U_{i1}^*, \dots, U_{in_m}^*$ become perfect predictors for choices as n grows large, whereas if $\tilde{\sigma}_n a(b_n) \to \infty$, the systematic parts become uninformative in the limit. For example, if $G(\eta)$ has very thin tails, the distribution of the maximum of J i.i.d. draws from $G(\cdot)$ becomes degenerate at a deterministic drifting sequence as J grows, and it is therefore necessary to increase the scale parameter σ in order for the scale of the maximum of idiosyncratic taste shifters to remain of the same order as differences in the systematic part. Specifically, if $\eta_{ij} \sim \Lambda(\eta)$, the extreme-value type I (or Gumbel) distribution, then we choose $b_n \asymp \frac{1}{2}\log n$ and $\sigma_n = 1$. For $\eta_{ij} \sim N(0, \sigma^2)$, it follows from known results from extreme value theory that the constants can be chosen as $b_n \asymp \sigma\sqrt{W\left(\frac{n}{2\pi}\right)} \asymp \sigma\sqrt{\log n}$ and $\sigma_n \asymp \frac{b_n}{\sigma^2}$, where $W(x)$ is the Lambert-W (product log) function, and for Gamma-distributed η_{ij}, $b_n \asymp \log n$ and $\sigma_n = 1$.[5]

The remainder of this section gives the main convergence result in Theorem 4.2 and Corollary 4.1, and a qualitative outline of the main technical arguments. For the first step, Lemma B.1 in the appendix establishes that if the distribution of η is in the domain of attraction of the extreme-value type-I distribution, then for large sets of alternatives, the implied conditional choice probabilities can be approximated by those implied by the Logit model. For expositional clarity, the main text only states the main approximation result in terms of the "unilateral" decision problem of a single agent facing a choice over an increasing set of alternatives.

[4]Note also that choosing $\tilde{J}_n = [\alpha n^{1/2}]$ for a given choice of $\alpha > 0$ would be equivalent to the original rate $J_n = [\sqrt{n}]$ with a different value for the intercept of the random utility from the outside option, so that our implicit choice $\alpha = 1$ is only a normalization.

[5]See Resnick (1987), pages 71-72.

Theorem 4.1. *Suppose that Assumptions 2.1, 2.2, and 4.1 hold, and that $z_1, \dots, z_J$ are J i.i.d draws from a distribution $M(z)$ with p.d.f. $m(z)$. Then as $J \to \infty$,*

$$\begin{aligned} P(U_{i0} \geq U_{ik}, k = 0, \dots, J) &\to \frac{1}{1 + \int \exp\{U(x_i, s)\} m(s) ds} \\ JP(U_{ij} \geq U_{ik}, k = 0, \dots, J | z_j = z) &\to \frac{\exp\{U(x_i, z)\}}{1 + \int \exp\{U(x_i, s)\} m(s) ds} \end{aligned}$$

almost surely for any fixed $j = 1, 2, \dots, J$.

See the appendix for a proof. Note that the limits on the right-hand side correspond to the choice probabilities for random sets of alternatives under the IIA assumption which were derived by Dagsvik (1994). It is important to note that the rate of convergence to the limiting choice probabilities depends crucially on the shape of the tails of $G(s)$. While for some choices for $G(s)$, convergence may be very fast, in the case of the standard normal distribution the rate of approximation for the c.d.f. of the maximum is as slow at $1/\log n$ (see e.g. Hall (1979)). Hence, Theorem 4.1 suggests that CCPs resulting from extreme-value type-I taste shifters can be viewed as a reference case for modeling choice among large sets of alternatives even if convergence to that limit may be very slow for alternative specifications. The extreme value approximation to conditional choice probabilities also requires that individuals' opportunity sets grow in size as the market gets large. We show in Lemma B.2 in the appendix that the opportunity sets W_i and M_j do indeed grow at the rate $n^{1/2}$ with common bounds that hold for all individuals simultaneously with probability approaching one.

For the second step, we need to address that woman i's opportunity set is in general endogenous with respect to her own taste shifters. To this end, Lemma B.3 in the appendix establishes that the conditional distribution of taste shifters given the opportunity sets converges to its marginal distribution. It is important to point out that convergence does not depend on any assumptions regarding how the data generating process selects among the multiple stable matchings. Instead, our argument is based on bounds on inclusive values based on the M- and W-preferred stable matchings, respectively.

For the third step, we show that woman i's conditional choice probabilities given her opportunity set can be approximated using a state variable $\hat{\Gamma}_w(x_i)$ that depends only on her observable characteristics x_i. Specifically, we consider the inclusive values associated with the extremal matchings, where for the M-preferred matching, we denote

$$\begin{aligned} I_{wi}^M &:= I_{wi}[M_i^M] = \frac{1}{J} \sum_{j \in M_i^M} \exp\{U(x_i, z_j)\} \\ I_{mj}^M &:= I_{mj}[W_j^M] = \frac{1}{J} \sum_{i \in W_j^M} \exp\{V(z_j, x_i)\} \end{aligned}$$

and we also write I_{wi}^W and I_{mj}^W for the inclusive values resulting from the W-preferred matching, and I_{wi}^* and I_{mj}^* for any other stable matching. In analogy to our definitions given exogenously fixed choice sets, we define the average inclusive value function under the M-preferred matching as

$$\hat{\Gamma}_w^M(x) := \frac{1}{J^2}\sum_{j=1}^{n_m}\frac{\exp\{U(x,z_j)+V(z_j,x)\}}{1+I_{mj}^M}$$
$$\hat{\Gamma}_m^M(z) := \frac{1}{J^2}\sum_{i=1}^{n_w}\frac{\exp\{U(x_i,z)+V(z,x_i)\}}{1+I_{wi}^M}$$

and, we define $\hat{\Gamma}_w^W(x)$, $\hat{\Gamma}_m^W(z)$, $\hat{\Gamma}_w^*(x)$, and $\hat{\Gamma}_m^*(z)$ in a similar manner for the W-preferred, or some generic matching, respectively. Since the opportunity sets M_i^* and W_j^* arising from any stable matching satisfy $M_i^M \subset M_i^* \subset M_i^W$ and $W_j^W \subset W_j^* \subset W_j^M$, we immediately obtain the relations

$$I_{wi}^M \le I_{wi}^* \le I_{wi}^W, \text{ and } I_{mj}^M \ge I_{mj}^* \ge I_{mj}^W$$

and

$$\hat{\Gamma}_w^M(x) \le \hat{\Gamma}_w^*(x) \le \hat{\Gamma}_w^W(x), \text{ and } \hat{\Gamma}_m^M(z) \ge \hat{\Gamma}_m^*(z) \ge \hat{\Gamma}_m^W(z)$$

for all values of x and z, respectively. Hence, we can use the average inclusive value functions corresponding to the extremal matchings to bound those associated with any other stable matching.

Since $U(x,z)$ is bounded, we obtain

$$I_{wi}^* - \hat{\Gamma}_w^*(x_i) = \frac{1}{n^{1/2}}\sum_{j=1}^{n_m}\exp\{U(x_i,z_j)\}\left[\mathbb{1}\{V_{ji}\ge V_j^*[W_j^*]\} - \Lambda_m(z_j,x_i;W_j^*)\right] + o_p(1)$$

Then, Lemma B.3 also implies that the random variables

$$v_{ji}(I_{mj}^*) := \mathbb{1}\{V_{ji}\ge V_j^*[W_j^*]\} - \Lambda_m(z_j,x_i;I_{mj}^*)$$

are approximately independent across $j=1,\dots,n_m$ conditional on I_{mj}^*. Hence, the difference $I_{wi}^*-\hat{\Gamma}_w^*(x_i)$ can be approximated as a weighted average of mean-zero random variables, where the pairwise correlations vanish sufficiently fast as n grows. This argument is made precise in Lemma B.4 in the appendix. It follows that

$$I_{wi}^* = \hat{\Gamma}_w^*(x_i) + o_p(1) \quad \text{ and } I_{mj}^* = \hat{\Gamma}_m^*(z_j) + o_p(1)$$

which allows us to approximate inclusive values as a function of observable characteristics alone.

We then proceed to step four and derive an (approximate) fixed point representation for the inclusive values. Note that by convergence of I_{mj}^* to $\hat{\Gamma}_m^*(z_j)$ and the continuous mapping theorem we can write

$$\hat{\Gamma}_w^*(x_i) = \frac{1}{n}\sum_{j=1}^{n_m} \frac{\exp\{U(x_i,z_j)+V(z_j,x_i)\}}{1+\hat{\Gamma}_m^*(z_j)} + o_p(1)$$

We next define the fixed point mapping

$$\begin{aligned} \hat{\Psi}_w[\Gamma_m](x) &= \frac{1}{n}\sum_{j=1}^{n_m} \frac{\exp\{U(x,z_j)+V(z_j,x)\}}{1+\Gamma_m(z_j)} \\ \hat{\Psi}_m[\Gamma_w](z) &= \frac{1}{n}\sum_{i=1}^{n_w} \frac{\exp\{U(x_i,z)+V(z,x_i)\}}{1+\Gamma_w(x_i)} \end{aligned} \tag{4.1}$$

In that notation, we can characterize the pairwise stability conditions as the fixed-point problem

$$\hat{\Gamma}_m^* = \hat{\Psi}_m[\hat{\Gamma}_w^*] + o_p(1) \quad \text{and } \hat{\Gamma}_w^* = \hat{\Psi}_w[\hat{\Gamma}_m^*] + o_p(1) \tag{4.2}$$

where, noting that $\hat{\Gamma}_m^*, \hat{\Gamma}_w^* \geq 0$ a.s., the remainder converges in probability to zero uniformly in Γ_w, Γ_m by Lemma B.4 and the continuous mapping theorem. In particular, the inclusive value functions for the two extremal matchings, $\hat{\Gamma}^M$ and $\hat{\Gamma}^W$ are solutions to the same fixed point problem.

We can now show that the solutions to the fixed-point problem in a finite economy (4.2) converge to the (unique) fixed point of the limiting problem (3.5): For one, the fixed-point mapping $\hat{\Psi}_n$ is a sample average over functions of x, z and Γ, and can be shown to converge in probability to its population expectation Ψ, which defines the limiting fixed point problem in (3.5). We then use this result to show that if a inclusive value function results from a stable matching, it can be represented as an approximate solution to the fixed point problem. Since the approximation is only shown to be valid for opportunity sets satisfying pairwise stability, the converse need not hold. However, the solution to the fixed point problem was shown to be unique in the previous section, and a stable matching is always guaranteed to exist, so that in fact the fixed point representation and pairwise stability are asymptotically equivalent.

Finally, since $\log \Psi[\cdot]$ was shown to be a contraction, the solution to (3.5) is unique and well-separated in the sense that large perturbations of Γ relative to the fixed point Γ^* also lead to sufficiently large changes in $\Gamma - \Psi[\Gamma]$. In particular, the functions $\hat{\Gamma}_w^M, \hat{\Gamma}_m^M$ and $\hat{\Gamma}_w^W, \hat{\Gamma}_m^W$ corresponding to the extremal matchings are solutions to (3.5) and therefore coincide even in the finite economy. Since the average inclusive value functions from the extremal matchings bound $\hat{\Gamma}_w^*, \hat{\Gamma}_m^*$ for any other stable matching, the functions $\hat{\Gamma}_w^*(x), \hat{\Gamma}_m^*(z)$ are uniquely determined.

Formally, we have the following theorem, which is proven in the appendix:

Theorem 4.2. *Suppose Assumptions 2.1, 2.2, and 4.1 hold. Then (a) for any stable matching, the inclusive values satisfy the fixed-point characterization in equation (4.2), and (b) we have convergence of the inclusive value functions* $\|\hat{\Gamma}_m^* - \Gamma^*\|_\infty \xrightarrow{p} 0$ *and* $\|\hat{\Gamma}_m^* - \Gamma^*\|_\infty \xrightarrow{p} 0$.

The preceding steps established that for a large number of market participants, we can find an approximate parametrization of the distribution of matched characteristics with the inclusive value functions Γ_w and Γ_m, which are characterized as solutions of a system of equilibrium conditions. The second part of Theorem 4.2 implies that the solutions of the sample equilibrium conditions converge to those of the limiting game discussed in the previous section. A converse of part (a) is not needed for our arguments, but with little additional work, it can be shown that it follows from existence of stable matchings in the finite economy together with uniqueness of the limit values of Γ established in Theorem 3.1.

Since the inclusive value functions are asymptotically sufficient for describing the distribution of matched characteristics, convergence of $\hat{\Gamma}$ to Γ_0 also implies convergence of the matching frequencies in the finite economy to the (unique) limit $f(x, z)$. Formally, we can now combine the conclusion of Theorem 4.2 with Lemma B.4 to obtain the limiting distribution:

Corollary 4.1. *For a given matching* μ *define the empirical matching frequencies* $\hat{F}_n(x, z|\mu) := \frac{1}{n}\sum_{i=0}^{n_w}\sum_{j\in\mu_m(i)} \mathbb{1}\{X_i \le x, Z_j \le z\}$, *where* $\mu_m(0)$ *denotes the set of men that are single under the matching* μ. *Then under the assumptions of Theorem 4.2, for any sequence of stable matchings* μ_n^*, $\hat{F}_n(x, z|\mu_n^*)$ *converges to a measure* F *with density* $f(x, z)$ *given in equation (3.6).*

Broadly speaking, this result states that the matched characteristics from any stable matching converge "in distribution" to the limiting model in section 3, where the measure F and the empirical matching frequencies have the properties of proper probability distributions except that their overall mass is not normalized to one. Note also that the empirical matching frequencies could equivalently be defined as $\hat{F}_n(x, z|\mu) := \frac{1}{n}\sum_{i=0}^{n_m}\sum_{i\in\mu_w(j)} \mathbb{1}\{X_i \le x, Z_j \le z\}$ where by the definition of a matching, $\mu_m(j)$ and $\mu_w(i)$ are singleton for all values of i, j except zero. This corollary is the main practical implication of our asymptotic analysis, and in the remainder of the paper we discuss some implications for identification and estimation of preference parameters from the distribution of matched characteristics.

5. Identification and Estimation

Since the main objective of this paper is to find estimators or tests that are consistent as the number of agents in the market increases, we analyze parameter identification from the limiting distribution of the matching game rather than its finite-market version.[6] For now,

[6]Specifically, our proofs of large-sample results for estimators are analogous to generic convergence arguments for extremum estimation (see e.g. Newey and McFadden (1994) sections 2 and 3, or van der Vaart and Wellner

this part of the paper assumes that K is small relative to the size of the market, $n_m + n_w$. In future work, we plan to extend our analysis to the case in which the sample includes a non-negligible fraction of the market.

5.1. **Matching Frequencies and Surplus.** We now show how to transform the equilibrium matching probabilities derived in section 3 into a demand system that allows us to analyze identification in a fairly straightforward manner. Recall that for the limiting game, the joint density of observable characteristics for a matched pair was given in (3.6) by

$$f(x,z) = \frac{\exp\{U(x,z) + V(z,x) + \gamma_w + \gamma_m\}w(x)m(z)}{(1+\Gamma_w^*(x))(1+\Gamma_m^*(z))} \qquad x \in \mathcal{X}, z \in \mathcal{Z}$$

where Γ_w^* and Γ_m^* solve the fixed point problem

$$\begin{aligned}\Gamma_w^*(x) &:= \int \frac{\exp\{U(x,s) + V(s,x) + \gamma_m\}m(s)}{1+\Gamma_m^*(s)}ds \\ \Gamma_m^*(z) &:= \int \frac{\exp\{U(s,z) + V(z,s) + \gamma_w\}w(s)}{1+\Gamma_w^*(s)}ds\end{aligned}$$

We also define the marital *pseudo-surplus* of a match as the sum of the deterministic parts of random payoffs,

$$W(x,z) := U(x,z) + V(z,x)$$

where in the absence of a common numeraire, the relative scales of men and women's preference are normalized by a multiple of the conditional standard deviation of random utilities given $x_i = x$ and $z_j = z$. In order to understand identifying properties of the model, it is important to note that we can express the distribution in terms of $W(x,z)$ alone:

$$f(x,z) = \frac{\exp\{W(x,z) + \gamma_w + \gamma_m\}w(x)m(z)}{(1+\Gamma_w^*(x))(1+\Gamma_m^*(z))} \qquad x \in \mathcal{X}, z \in \mathcal{Z} \tag{5.1}$$

Also, the fixed-point equations defining Γ_w^* and Γ_m^* can be rewritten as

$$\begin{aligned}\Gamma_w^*(x) &:= \int \frac{\exp\{W(x,s) + \gamma_m\}m(s)}{1+\Gamma_m^*(s)}ds \\ \Gamma_m^*(z) &:= \int \frac{\exp\{W(s,z) + \gamma_w\}w(s)}{1+\Gamma_w^*(s)}ds \end{aligned} \tag{5.2}$$

Hence, the joint distribution of matching characteristics depends on the systematic parts of U_{ij} and V_{ji} only through $W(x,z)$, so that we can in general not identify $U(x,z)$ and $V(z,x)$ separately without additional restrictions. This naturally limits the possibilities for welfare

(1996) chapter 3.2-3), where the population parameter of interest is the unique maximizer of the (limiting) population objective function, which is then approximated by sample quantities. These arguments do not require that the sampling objective function results from the same data generating process (DGP) as the limiting objective as long as we have uniform convergence along the triangular sequence of DGPs.

assessments from the perspective of either side of the market, but conversely, the pseudo-surplus $W(x,z)$ is, at least in the limit, a sufficient statistic for the impact of observable characteristics on the resulting stable matching.

5.2. **Sampling Distribution.** For identification and estimation, we assume that we observe a sample of K couples ("households") $k = 1, \dots, K$, where $w(k)$ and $m(k)$ give the respective indices of the wife and the husband, where $m(k) = 0$ if the kth unit represents of a single woman and $w(k) = 0$ if we observe a single man. We can now use the asymptotic measure of matched characteristics in (3.6) to obtain the density $h(x,z)$ of the sampling distribution of $(x_{w(k)}, z_{m(k)})$ depending on the sampling protocol.

First, consider the case of a random sample of individuals, where men and women are surveyed with the same probability. The resulting sample reports the spouses' characteristics $(x_{w(k)}, z_{m(k)})$ for the kth unit. If the selected individual is single, we only observe his own characteristics, and the spousal characteristics are coded as missing. Then the sampling distribution is given by the p.d.f.

$$\begin{aligned} h_1(x,z) &= \frac{2f(x,z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}} \\ h_1(x,*) &= \frac{f(x,*)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}} \\ h_1(*,z) &= \frac{f(*,z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}} \end{aligned} \tag{5.3}$$

Alternatively, if the survey design draws at random from the population couples ("households"), including singles, then the sampling distribution is characterized by the p.d.f.

$$\begin{aligned} h_2(x,z) &= \frac{f(x,z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\} - \int f(s,t)dtds} \\ h_2(x,*) &= \frac{f(x,*)}{\exp\{\gamma_w\} + \exp\{\gamma_m\} - \int f(s,t)dtds} \\ h_2(*,z) &= \frac{f(*,z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\} - \int f(s,t)dtds} \end{aligned}$$

This discussion could easily be extended to cases in which a survey contains sampling weights that allow us to reconstruct an weighted sample with comparable properties.[7] However, it is important that for our analysis the matched pair is not the unit of observation, but endogenous to the model. Since our identification results are based on likelihood ratios

[7]E.g. suppose that a survey samples a female with attributes x_i with probability $q_w(x_i)$, and a male with attributes z_j with probability $q_m(z_j)$, and that any sampled individual also reports the characteristics of their spouse. It can then easily be verified that the correct sampling weights would be $\frac{1}{q_w(x_i)}$ and $\frac{1}{q_m(z_j)}$, respectively, if i and j are single, and $\frac{1}{q_w(x_i)+q_m(z_j)}$ if they are married to each other.

rather than absolute levels, we discuss identification based on direct knowledge of the measure $f(x,z)$ rather than the sampling distribution.

5.3. **Identification.** We next consider identification of the pseudo-surplus function $W(x,z)$: Taking logs on both sides and rearranging terms, we obtain

$$\log f(x,z) - [\log w(x) + \log m(z)] = W(x,z) - [\log(1+\Gamma_w^*(x)) + \log(1+\Gamma_m^*(z))] \quad (5.4)$$

Note that the terms on the left-hand side are features of the distribution of observable characteristics and can be estimated from the data. This expression also suggests a computational shortcut which obviates the need to infer the (unobserved) opportunity sets $m(z,x)$ and $w(x,z)$ to compute the matching probabilities. We can use differencing arguments to eliminate the inclusive values $\Gamma_w^*(x)$ and $\Gamma_m^*(z)$ from the right-hand side expression in (5.4) using information on the shares of unmatched individuals. Specifically, for any values $x \in \mathcal{X}$ and $z \in \mathcal{Z}$, consider differences of the form

$$\log \frac{f(x,z)}{f(x,*)f(*,z)} = W(x,z)$$

Since the quantities on the right-hand side are observed, the joint distribution of matched characteristics identifies the pseudo-surplus function $W(x,z)$. This means we can directly estimate the strength of complementarities between observable types from differences in logs of the p.d.f. of characteristics in matched pairs.

This simplification can be seen as a direct consequence of the IIA assumption, where dependence of choice probabilities on a latent opportunity set is entirely captured by the inclusive value. Similar differencing arguments have been used widely in the context of discrete choice models for product demand, see Berry (1994).

Furthermore, we can identify the average inclusive value function from the conditional probabilities of remaining single given different values of x,

$$\Gamma_w(x) = \frac{w(x)\exp\{\gamma_w\}}{f(x,*)} - 1$$

We can summarize these findings in the following proposition:

Proposition 5.1. ***(Identification)*** *(a) The surplus function $W(x,z)$ and the inclusive value functions $\Gamma_w(x)$ and $\Gamma_m(z)$ are point-identified from the limiting measure $f(x,z)$. (b) Without further restrictions, the systematic parts of the random utilities, $U(x,z)$ and $V(z,x)$ are not separately identified from the limiting measure $f(x,z)$.*

While non-identification of the random utility functions $U(x,z)$ and $V(z,x)$ is a negative result, it is important to note that the surplus function $W(x,z) := U(x,z) + V(z,x)$ is an object of interest in itself. Most importantly, the characterization of the limiting distribution implies that knowledge of $W(x,z)$ is sufficient to compute any counterfactual distributions of match characteristics and analyze welfare consequences of policy changes. E.g. it is possible to predict the effect of changes in sex ratios on marriage rates based on the surplus function $W(x,z)$ alone.

However, in the context of a parametric model for $U(x,z;\theta)$ and $V(z,x;\theta)$ it may be possible to identify preference parameters for the two sides of the market separately in the presence of exclusion restrictions. Since in a two-sided matching market "demand" of one side constitutes "supply" on the other, this identification problem illustrates the close parallels with estimation of supply and demand functions from market outcomes. In that event, our asymptotic characterization of matching probabilities allows to formulate rank conditions for identification directly in terms of the surplus function $W(x,z)$. For the transferable utility case, Galichon and Salanié (2010) discuss identification for several specific models with restrictions of this type.

Comparing this result to the previous literature on identification in matching markets, it can be seen that the decision whether to model idiosyncratic preferences over types or individuals has crucial implications as to whether we can only identify marital surplus or individual utilities. Our identification result suggests that for "thick" markets, models with and without transferable utilities may lead to similar qualitative implications, however in the latter case information on actual transfers may be useful for identification.

5.4. **Welfare Effects.** In addition to prediction of counterfactuals regarding observable characteristics of the realized matches, our model also allows for welfare evaluations of policy interventions. Recall that the inclusive value is related to the expectation of indirect utility via

$$\mathbb{E}[U_i^*|X_i=x] = \log(1+\Gamma_w(x)) + \kappa$$

where $\kappa \approx 0.5772$ is Euler's constant. In particular, if we define $s_w(x) := \frac{f(x,*)}{w(x)\exp\{\gamma_w\}}$ as the share of women of type x that remain single, we can express the inclusive value in terms of observable quantities,

$$\mathbb{E}[U_i^*|X_i=x] = -\log\frac{f(x,*)}{w(x)} + \gamma_w + \kappa = -\log s_w(x) + const$$

Hence, for natural experiments that change the composition of matching markets, we can interpret the difference in the log shares of unmatched individuals directly as the average change in the surplus from participating in the matching market for individuals of a given observable type.

For example we can evaluate the effect of changes to the marginal distribution of characteristics in the market, $w(x)$ and $m(z)$ on individual surplus for any type on either side of the market. If observed characteristics include income and own education, we can identify the monetary return to education on the marriage market (compensating or equivalent variation) from local shifts in education levels and incomes that leave the share of unmatched individuals constant.[8] Clearly, changes in individuals' types also affect the value of their outside option - which was normalized to zero in our analysis - so that comparisons based $\log s_w(x)$ do not capture welfare changes that do not operate through the matching market. For example, we may observe that an increase in women's education leads to an increase in the share of unmarried women. This may either reflect a deterioration of women's matching prospects, or an increase in the relative value of their outside option, either of which is associated with a decrease in her expected (net) return from participating in the matching market.

5.5. **Estimation.** We now turn to estimation of a parametric model for the systematic payoffs $U^*(x,z) := U^*(x,z;\theta)$ and $V^*(z,x) := V^*(z,x;\theta)$, where the unknown parameter $\theta \in \Theta_0$, a compact subset of some Euclidean space. We can then estimate the parameter θ up to the normalizations necessary for identification using the resulting surplus function $W(x,z) = W(x,z;\theta) := U^*(x,z;\theta) + V^*(z,x;\theta)$ Without loss of generality, assume that estimation is based on a random sample of individuals rather than couples, with sampling distribution given in (5.3). Considering log-ratios of the density of match characteristics is convenient for the analysis of identification of the model without explicitly characterizing the (unobserved) set of available spouses for any individual. However, two-step estimation of payoff parameters based on first-step estimates of the p.d.f. of the sampling distribution $h_1(x,z)$ is impractical for estimation, since fully nonparametric estimation of the density would suffer from a curse of dimensionality in most realistic settings. On the other hand, imposing the restrictions resulting from our knowledge on the function $W(x,z)$ when estimating $h_1(x,z)$ requires solving for the equilibrium distribution and latent inclusive values, which we wanted to avoid in the first place.

An alternative consists in treating the inclusive values $\Gamma_{ww(k)} := \Gamma_w(x_{w(k)})$, $\Gamma_{mm(k)} := \Gamma_m(z_{m(k)})$ as auxiliary parameters in maximum likelihood estimation of parameters of the surplus function $W(x,z)$, and imposing equilibrium conditions as side constraints in a joint maximization problem over θ and $\Gamma_{ww(k)}, \Gamma_{mm(k)}$, where $k = 1, \dots, K$. Specifically, it follows from equation (5.3) that

$$
\begin{aligned}
l_K(\theta,\Gamma) &:= \log h_1(x_{w(k)}, z_{m(k)}|\theta,\Gamma) \\
&= W(x_{w(k)}, z_{m(k)};\theta) + \log(2)\mathbb{1}\{w(k) \neq 0, m(k) \neq 0\} \\
&\quad - \log(1+\Gamma_w(x_{w(k)})) - \log(1+\Gamma_m(z_{m(k)})) + \text{const}
\end{aligned}
$$

[8]see Small and Rosen (1981)

Hence we can write the likelihood function (up to a constant) as

$$\begin{aligned} L_K(\theta,\Gamma) &:= \sum_{k=1}^{K}\Big(W(x_{w(k)},z_{m(k)};\theta)+\log(2)\mathbb{1}\{w(k)\neq 0, m(k)\neq 0\} \\ &\quad -\log(1+\Gamma_w(x_{w(k)}))-\log(1+\Gamma_m(z_{m(k)}))\Big) \end{aligned}$$

We can also state the equilibrium conditions using the operator $\hat{\Psi}_K := (\hat{\Psi}_{wK}, \hat{\Psi}_{mK})$ where

$$\begin{aligned} \hat{\Psi}_{wK}[\Gamma](x) &= \frac{1}{K}\sum_{k=1}^{K}\frac{\exp\left\{W(x,z_{m(k)};\theta)\right\}\mathbb{1}\{m(k)\neq 0\}}{1+\Gamma_m(z_{m(k)})} \\ \hat{\Psi}_{mK}[\Gamma](z) &= \frac{1}{K}\sum_{k=1}^{K}\frac{\exp\left\{W(x_{w(k)},z;\theta)\right\}\mathbb{1}\{w(k)\neq 0\}}{1+\Gamma_w(x_{w(k)})} \end{aligned}$$

Hence the maximum likelihood estimator $\hat{\theta}$ solves

$$\max_{\theta,\Gamma} L_K(\theta,\Gamma) \quad \text{s.t.}\Gamma = \hat{\Psi}_K(\Gamma) \tag{5.5}$$

We can find the MLE $\hat{\theta}$ by constrained maximization of the log-likelihood, where treating the equilibrium conditions as constraints obviates the need of solving for an equilibrium in Γ at every maximization step. This constrained problem is high-dimensional in that Γ generally has to be evaluated at up to $2K$ different arguments, leading to $2K$ constraints. That system of constraints is not sparse but collinear, and we may adapt the constrained MPEC algorithm proposed and described in Dubé, Fox, and Su (2012) and Su and Judd (2012).

Since the equilibrium conditions for the inclusive value functions are market-specific, consistent estimation will generally require that we observe a large number of individuals or couples in each market. However in the absence of market-level heterogeneity, data from any finite number of markets is typically sufficient for consistency since all objects of interest are identified from the distribution in a single market. Asymptotic inference for surplus parameters or counterfactuals is standard if the sample used by the researcher contains a small fraction of the households or individuals in each market, and is drawn at random from the relevant population. If the researcher's sample includes the entire market, or a significant share of all individuals, then structural inference on surplus parameters has to account for conditional dependence of matching decisions across pairs. Asymptotics of this type have been developed in Menzel (2012) for discrete games, however an application of these ideas to matching markets is not straightforward and will be left for future research.

6. Matching with Limited Awareness

The asymptotic results on uniqueness of the limiting distribution and convergence in the previous sections require that the outside option remain relevant even as the number of available partners grows for each agent in the market. The asymptotic sequence of markets in our limiting experiment involved drawing an increasing number of random utilities for the outside option, and was motivated by approximating properties. A more plausible behavioral explanation why many agents remain unmatched even in large markets is that with a growing size of the market, each agent may not be aware of all her/his matching opportunities. Our asymptotic results require that the number of available matches grows for each agent, most importantly the step in Lemma B.4, and cannot be easily adjusted to a setup in which opportunity sets remain small.

This section develops an extension of our baseline setup in which every agent only becomes aware of a subset of potential matches at random, where the probability of meeting a potential spouse may be a function of observed characteristics. For example, characteristics may include geographic location, so that the probability of meeting may be a function of spatial distance. This modified thought experiment maintains that the number of draws for the outside option grow at a root-n rate with the size of the market and continues to assume that opportunity sets be large. The main purpose of this extension is to arrive at a more realistic interpretation of the pseudo-surplus function $W^*(x,z)$ in a setting in which agents may only observe a subset of agents in the market.

In the modified model, nature initially draws random utilities U_{ij}, V_{ji} according to the model in (2.1) and (2.2). The matching market then operates in two stages: in stage 1, agents meet at random and independently of the realized random matching payoffs, where the probability of a woman of type x meeting a man of type z is given by $r(x,z) \in [0,1]$. Awareness is assumed to be mutual, i.e. woman i is aware of man j if and only if man j is also aware of woman i. If a pair (i,j) of a woman and a man is not aware of each other, the random payoffs for a match between them are set to minus infinity, and otherwise equal to their initial values, i.e.

$$\tilde{U}_{ij,n} = \begin{cases} U_{ij,n} & \text{if } i \text{ and } j \text{ meet} \\ -\infty & \text{otherwise} \end{cases} \qquad \text{and } \tilde{V}_{ji,n} = \begin{cases} V_{ji,n} & \text{if } i \text{ and } j \text{ meet} \\ -\infty & \text{otherwise} \end{cases}$$

In stage 2, the market mechanism determines a matching that is stable with respect to the modified payoffs $\tilde{U}_{ij,n}, \tilde{V}_{ji,n}$, and which is observed by the researcher. Note that in the presence of an outside option of remaining single, the modified payoffs continue to satisfy the assumptions of Gale and Shapley (1962)'s model of stable marriage. In particular, our model of matching with limited awareness is guaranteed to produce at least one stable matching, and the set of stable matchings is a lattice under the preference orderings of the male and female side of the market, respectively.

It is then straightforward to verify that the argument leading to Corollary 4.1 can be adjusted to accommodate this extension, and we obtain the following limiting result:

Proposition 6.1. *Suppose Assumptions 2.1, 2.2, and 4.1 hold, and furthermore* $|\log r(x,z)| \leq \bar{R}$*, a finite constant, for all* $x \in \mathcal{X}$ *and* $z \in \mathcal{Z}$*. Then the distribution of matched characteristics is characterized by equations (5.1) and (5.2), where*

$$W^*(x,z) := U^*(x,z) + V^*(x,z) + \log r(x,z)$$

In particular, $W^*(x,z)$ *is nonparametrically identified from the measure* $f(x,z)$*.*

We can now directly apply our results on identification and estimation of the pseudo-surplus function to the modified matching game. Most importantly, we can see from the form of the limiting distribution of matchings in Proposition 6.1 that for large markets we cannot separate whether observable characteristics affect the matching frequency through the likelihood of meeting or preferences.

6.1. **Scaling Properties.** It is furthermore interesting to analyze the returns to scale with respect to market size on the matching function implied by our limiting model. For their model of transferable utilities with finite types, Choo and Siow (2006) show that the matching function is homogeneous of degree zero, so that in particular the rate of singles in the population does not depend on the number of market participants. However, our setup differs qualitatively from theirs in that we assume that idiosyncratic taste-shifters are individual- rather than type-specific, so that a larger number of matching opportunities increases the attractiveness of the best available partner relative to the outside option. Furthermore the limiting matching function in (3.6) corresponding to the nontransferable utility model is different from their result.

While under our asymptotics, all markets are assumed to grow large, our setup still allows for differences in relative scale by varying the parameters γ_m, γ_w. We can see that in general, the matching functions are not homogeneous of degree zero, but other things equal, increasing γ_m and γ_w simultaneously leads to a decrease in the share of singles. However, in the model with limited awareness we can eliminate scale effects by assuming that $\gamma_w + \log r(x,z)$ and $\gamma_m + \log r(x,z)$ are constant across markets.[9] We can interpret this comparison as individuals in each market being aware of roughly the same number of potential spouses regardless of market size.

Finally, we can analyze the effect of an increase of γ_w that leaves γ_m constant. From the fixed-point conditions (3.4) we can see that such a change would result in an increase of

[9]Note that under this restriction, the fixed-point conditions in (3.5) are solved by the same values for Γ_w and Γ_m for each market, so that the share of singles implied by (3.6) remains the same.

inclusive values for men and a decrease in inclusive values for women. From (3.6) we can then conclude that as a result, the share of singles among men would decrease, whereas the share of singles among women would increase. From an analogous argument, altering the marginal distributions $w(x)$ and $m(z)$ would result in similar effects on those types of men or women that become more "scarce" or "abundant," respectively.

7. Monte Carlo Simulations

In order to illustrate the different aspects of the theoretical convergence result, we simulate a very basic version of our model. We generate payoff matrices from the random utility model

$$\begin{aligned} U_{ij} &= U^*(x_i, z_j) + \sigma_n \eta_{ij} \\ V_{ji} &= V^*(z_j, x_i) + \sigma_n \zeta_{ji} \end{aligned}$$

where idiosyncratic taste shifters η_{ij}, ζ_{ji} are i.i.d. draws from a standard normal or extreme value type-I distribution. We then find the M- and W-preferred matchings using the Gale-Shapley (deferred acceptance) algorithm.

The first set of simulation experiments is meant to illustrate convergence of matching frequencies generated by the model to the limiting choice probabilities predicted by the asymptotic arguments in the previous sections. Specifically, we illustrate three qualitative conclusions of our theoretical results: (1) the degree of multiplicity of matching outcomes increases in the size of the market, but (2) that growth is not fast enough to affect the limiting distribution of matched characteristics. (3) Convergence of conditional choice probabilities to their extreme value limits when the respective distribution of η_{ij} and ζ_{ji} are not extreme-value type I can very slow in some cases, most importantly when taste shifters are generated from the standard normal distribution.

7.1. Approximation of Matching Probabilities.

7.1. Approximation of Matching Probabilities. Our first set of simulation results is based on a design with taste-shifters η_{ij}, ζ_{ji} generated from the extreme-value type I distribution and no observable characteristics. In table 7.1 we report the difference in the average size of a woman's opportunity set between the extremal matchings, $|M_i^W| - |M_i^M|$, where M_i^W denotes woman i's opportunity set under the female-preferred matching, and M_i^M her opportunity set under the male-preferred matching. As argued before, opportunity sets arising from stable matchings are nested: For any given stable matching μ^*, woman i's opportunity set under μ^* satisfies $M_i^M \subset M_i[\mu^*] \subset M_i^W$ with probability 1. Hence we can interpret the difference between the two extremal matchings as an upper bound on the variation of opportunity sets across different stable matchings. We also report the number of women for whom M_i differs across matchings, i.e. $M_i^W \not\subseteq M_i^M$ and the average inclusive values for the $W-$ and $M-$preferred matchings.

n	$\lvert M_i^W\rvert - \lvert M_i^M\rvert$	$\sharp\{M_i^W \neq M_i^M\}$	$I_w[M_i^W]$	$I_w[M_i^W] - I_w[M_i^M]$
10	0.0350	0.30	6.73	0.08
20	0.2125	3.05	7.50	0.35
50	0.1370	5.90	6.92	0.14
100	0.0905	8.35	6.77	0.07
200	0.1175	21.95	6.82	0.06
500	0.0574	27.50	6.87	0.02
1000	0.0539	52.00	6.85	0.01
2000	0.0510	92.60	6.86	0.01
5000	0.0041	102.00	6.87	0.00

Table 1. Comparisons of the Male and Female Preferred Matchings (Extreme-Value Type I Taste Shifters).

From table 7.1, we can see that as n grows, the number of women for whom there is a difference in opportunity sets across stable matchings increases steadily, although at a rate that is less than proportional to n. Furthermore, the average difference in the number of available spouses decreases in n, as does the difference in inclusive values. For the latter, this is in part due to the normalization of the inclusive values which are scaled by the inverse of root-n. Given that normalization, inclusive values converge to a nonzero limit, and the simulations also show that the variance of $I_w[M_i^*]$ (not reported in the table) decreases to zero. These simulation results suggest that realizations of payoffs that support an exponentially increasing number of stable matchings as in Theorem 3.19 in Roth and Sotomayor (1990) are not "typical" for the random utility model analyzed in this paper.

To verify the quality of the approximation to the distribution of matched types, we also compare the probability of remaining single predicted by the model and the simulated frequency of singles in the $W-$preferred stable matching. In this setting, the matching probabilities are the same under the $W-$ and $M-$preferred matching by the "Rural Hospital Theorem." The simulation results are reported in table 7.1. We can see that the quality of the approximation improves as n grows large, however markets have to be quite large for the approximation bias to be small, say less than half a percentage point. The asymptotic arguments suggest that the convergence rate for matching probabilities should be $n^{-1/4}$, regardless whether observed covariates are continuous or discrete.

Next, we repeat the same experiment with standard normal taste shifters to evaluate the quality of the extreme-value approximation for the conditional choice probabilities. In general, convergence rates for distributions of extremes depend on the shape of the tails of the c.d.f. of random taste shifters, and especially for a thin-tailed distribution the convergence rate is very slow.[10] Hence we should not expect the approximation to improve very much

[10]**?** showed that the rate of convergence for the c.d.f. of the maximum of independent normals is $\log n$.

n	Model Probability	Simulated Frequency	Bias (2 minus 4)
10	0.1265	0.1760	-0.0495
20	0.1265	0.1510	-0.0245
50	0.1265	0.1392	-0.0127
100	0.1265	0.1358	-0.0093
200	0.1265	0.1322	-0.0057
500	0.1265	0.1333	-0.0068
1000	0.1265	0.1338	-0.0074
2000	0.1265	0.1300	-0.0035
5000	0.1265	0.1270	-0.0005

Table 2. Theoretical and Simulated Matching Frequencies (Extreme-Value Type I Taste Shifters).

for moderate sample sizes. Table 7.1 reports results on differences in inclusive values and

n	$\lvert M_i^W\rvert - \lvert M_i^M\rvert$	$\sharp\{M_i^W \neq M_i^M\}$	$I_w[M_i^W]$	$I_w[M_i^W] - I_w[M_i^M]$
10	0.2340	1.86	7.37	0.55
20	0.2820	3.86	7.30	0.47
50	0.1920	7.90	7.20	0.20
100	0.1476	11.54	7.13	0.11
200	0.1196	21.18	7.21	0.06
500	0.0583	26.10	7.08	0.02
1000	0.0461	44.18	6.98	0.01

Table 3. Comparisons of the Male and Female Preferred Matchings (Standard Normal Taste Shifters).

opportunity sets across stable matchings, and it is interesting to see that the qualitative results on multiplicity of stable matchings remain unchanged as we alter the distribution of unobservables. Most importantly, the number of agents for whom there is a difference between the extremal matchings still grows in n, whereas the inclusive values converge to a common limit, independently of the chose stable matching. We also compare the predicted (asymptotic) probability of remaining single with the simulated frequencies in table 7.1. Here we can see that even for large markets, matching frequency approach the theoretical limit only up to a point, and then convergence becomes very slow. This is to be expected in light of the slow convergence rates of normal extremes, and should be read as a caveat on the approximations for distributions other than the normal. For realistic market sizes it would therefore be more plausible to impose extreme-value type I taste shifters as an assumption rather than arising from a many-alternative limit. However, the asymptotic

n	Model Probability	Simulated Frequency	Bias (2 minus 4)
10	0.1265	0.0880	0.0385
20	0.1265	0.0800	0.0465
50	0.1265	0.1040	0.0225
100	0.1265	0.1076	0.0189
200	0.1265	0.1052	0.0213
500	0.1265	0.1060	0.0205
1000	0.1265	0.1082	0.0183

Table 4. Theoretical and Simulated Matching Frequencies (Standard Normal Taste Shifters).

results on conditional choice probabilities still imply that the conditional Logit specification is the only asymptotically stable model for taste shifters that have distributions with tails of type I.

7.2. **ML Estimation of Preference Parameters.** Finally we give some Mote Carlo evidence on estimation of structural preference parameters from the realized match. Our simulation design specifies systematic utilities

$$U^*(x_i, z_j) = \alpha + \beta x_i + \delta x_i z_j$$
$$V^*(z_j, x_i) = \alpha + \beta z_j + \delta x_i z_j$$

where types for men and women, $x_i, z_j \in \{0, 1\}$ are generated from a symmetric Bernoulli distribution, and idiosyncratic taste shifters η_{ij}, ζ_{ji} are i.i.d. draws from the extreme-value type I distribution. The parameter β measures the systematic difference in tastes for marriage between the two types, and $\delta \neq 0$ generates matchings that are assortative across observable type categories.

Since with discrete types, the limiting estimation problem is finite-dimensional, we use the fmincon command in Matlab to obtain the constrained maximum likelihood estimators. Computation using the standard algorithm is fast and stable. Table 7.2 reports the median and normalized interquartile range[11] as robust estimates of location and scale of the distribution of the estimators to accommodate outliers due to numerical problems for small sample sizes.

We can see that for matching markets of moderate to large sizes, the estimates become concentrated near the values specified in the data generating process, and the standard

[11]We multiply the interquartile range by 0.7413 for the estimate to match the standard deviation in the case of a normal distribution.

n	α		β		δ	
50	0.3937	(0.3934)	0.5944	(0.7544)	0.9384	(0.2359)
100	0.4802	(0.4795)	0.4648	(0.5529)	0.9098	(0.2145)
200	0.4511	(0.1545)	0.4879	(0.4078)	0.9980	(0.1288)
500	0.5105	(0.1329)	0.4929	(0.2315)	0.9630	(0.0700)
1000	0.4547	(0.1199)	0.4715	(0.1952)	0.9884	(0.0788)
(DGP)	0.50		0.50		1.00	

TABLE 5. Distribution of Maximum Likelihood Estimators for Preference Parameters: Median and normalized Interquartile Range (in parentheses)

deviation of the estimator decreases as the market size increases. A future version of this paper will include simulation results for larger markets.

APPENDIX A. PROOFS FOR RESULTS FROM SECTION 3

A.1. **Proof of Theorem 3.1:** Without loss of generality, let $\gamma^* = 0$. The proof consists of three steps: We first show that under Assumption 2.1, any solution to the fixed point problem in (3.5) is differentiable, so that we can restrict the problem to fixed points in a Banach space of continuous functions. We then show that the mapping $(\log \Gamma_w, \log \Gamma_m) \mapsto (\log \Psi_w[\Gamma_m], \log \Psi_m[\Gamma_w])$ is a contraction, so that the conclusions of the theorem follow from Banach's fixed point theorem. Without loss of generality, we only consider the case in which all observable characteristics are continuously distributed, $x_{1i} = x_i$ and $z_{1i} = z_i$.

Bounds on solutions. We first establish that any pair of functions $(\Gamma_w^*(x), \Gamma_m^*(z))$ solving the fixed point problem in (3.5) are bounded from above: Assuming the solutions exist, and noticing that $\Gamma_m(z) \geq 0$ for all $z \in \mathcal{Z}$, we have that

$$\begin{aligned} \Gamma_w^*(x) &= \Psi_w[\Gamma_m^*](x) = \int \frac{\exp\{U(x,s) + V(s,x)\} m(s)}{1 + \Gamma_m^*(s)} ds \\ &\leq \int \exp\{U(x,s) + V(s,x)\} m(s) ds \leq \exp\{\bar{U} + \bar{V}\} \end{aligned} \tag{A.1}$$

which is finite by assumption 2.1. Similarly, we can see that

$$\Gamma_m^*(z) \leq \exp\{\bar{U} + \bar{V}\} \tag{A.2}$$

if a solution to the fixed point problem exists.

Continuity of solutions. In order to establish continuity, suppose that (Γ_w, Γ_m) is a fixed point of (Ψ_w, Ψ_m), in particular Γ_m has to satisfy

$$\Gamma_w = \Psi_w[\Psi_m[\Gamma_w]]$$

Hence, consecutive application of Ψ_w and Ψ_m gives

$$\Psi_w[\Psi_m[\Gamma_w]](x) = \int \frac{\exp\{U(x,t) + V(t,x)\}}{1 + \int \frac{\exp\{U(s,z)+V(z,s)\} w(s)}{1+\Gamma_w(s)} ds} m(t) dt$$

Since $\exp\{U(x,z)\}$ and $\exp\{V(z,x)\}$ are also continuous in z, x, and the integrals are all nonnegative, $\Psi_w[\Psi_m[\Gamma_w]]$ is also bounded and continuous in x for any nonnegative function Γ_w. Similarly, $\Psi_m[\Psi_w[\Gamma_m]]$

is also bounded and continuous, so that any solution of the fixed point problem in (3.5), if one exists, must be continuous. Existence of bounded derivatives up to the pth order follows by induction using the product rule and existence of bounded partial derivatives of the functions $U(x,z)$ and $V(z,x)$, see Assumption 2.1.

Hence, the range of the operators $\Psi_w \circ \Psi_m$ and $\Psi_m \circ \Psi_w$ is restricted to a set of bounded continuous functions, so that we can w.l.o.g. restrict the fixed point problem to the space of continuous functions satisfying the bounds derived before.

Contraction mapping: We next show that the mapping $(\log \Gamma_w, \log \Gamma_m) \mapsto (\log \Psi_w[\Gamma_m], \log \Psi_m[\Gamma_w])$ is a contraction on a Banach space of functions that includes all potential solutions of the fixed point problem (3.5). Specifically, let $\mathcal{C}^*$ denote the space of continuous functions on $\mathcal{X} \times \mathcal{Z}$ taking nonnegative values and satisfying (A.2) and (A.1). As shown above, any solution to the fixed point problem - if a solution exists - is an element of $\mathcal{C}^* \times \mathcal{C}^*$, which is a Banach space.

Consider alternative pairs of functions (Γ_w, Γ_m) and $(\tilde{\Gamma}_w, \tilde{\Gamma}_m)$. Using the definitions of the operators,

$$\begin{aligned}\log \Psi_w[\tilde{\Gamma}_m](x) - \log \Psi_w[\Gamma_m](x) &= \log \int \frac{\exp\{U(x,s)+V(s,x)\}m(s)}{1+\exp\left\{\log \tilde{\Gamma}_m(s)\right\}} ds \\ &\quad - \log \int \frac{\exp\{U(x,s)+V(s,x)\}m(s)}{1+\exp\{\log \Gamma_m(s)\}} ds\end{aligned}$$

By the mean-value theorem for real-valued functions of a scalar variable, for every value of x, there exists $t(x) \in [0,1]$ such that

$$\begin{aligned}\log \frac{\Psi_w[\tilde{\Gamma}_m](x)}{\Psi_w[\Gamma_m](x)} &= -\frac{1}{\Psi_w\left[\Gamma_m^{1-t(x)}\tilde{\Gamma}_m^{t(x)}\right](x)} \\ &\quad \times \int \frac{\exp\{U(x,s)+V(s,x)\}\Gamma_m(s)^{1-t(x)}\tilde{\Gamma}_m(s)^{t(x)}}{\left[1+\Gamma_m(s)^{1-t(x)}\tilde{\Gamma}_m(s)^{t(x)}\right]^2}\left[\log \tilde{\Gamma}_m(s) - \log \Gamma_m(s)\right] m(s) ds\end{aligned}$$

pointwise in x. Since we are restricting our attention to functions $\Gamma_m(z), \tilde{\Gamma}_m(z)$ satisfying the bounds in equation (A.1), we can bound the ratio

$$0 \le \frac{\Gamma_m(z)^{1-t(x)}\tilde{\Gamma}_m(z)^{t(x)}}{1+\Gamma_m(z)^{1-t(x)}\tilde{\Gamma}_m(z)^{t(x)}} \le \frac{\exp\{\bar{U}+\bar{V}\}}{1+\exp\{\bar{U}+\bar{V}\}} =: \lambda \tag{A.3}$$

for all $z \in \mathcal{Z}$. Since all components of the integrand are nonnegative we can bound the right hand side in absolute value by

$$\begin{aligned}\left|\log \frac{\Psi_w[\tilde{\Gamma}_m](x)}{\Psi_w[\Gamma_m](x)}\right| &\le \frac{\lambda}{\Psi_w\left[\Gamma_m^{1-t(x)}\tilde{\Gamma}_m^{t(x)}\right](x)} \int \frac{\exp\{U(x,s)+V(s,x)\}}{1+\Gamma_m(s)^{1-t(x)}\tilde{\Gamma}_m(s)^{t(x)}} \sup_{z\in\mathcal{Z}}\left|\log \tilde{\Gamma}_m(s) - \log \Gamma_m(s)\right| m(s) ds \\ &= \frac{\lambda}{\Psi_w\left[\Gamma_m^{1-t(x)}\tilde{\Gamma}_m^{t(x)}\right](x)} \left\|\log \tilde{\Gamma}_m - \log \Gamma_m\right\|_\infty \int \frac{\exp\{U(x,s)+V(s,x)\}}{1+\Gamma_m(s)^{1-t(x)}\tilde{\Gamma}_m(s)^{t(x)}} m(s) ds \\ &= \lambda \left\|\log \tilde{\Gamma}_m - \log \Gamma_m\right\|_\infty\end{aligned}$$

since the integral in the second to last line is equal to $\Psi_w\left[\Gamma_m^{1-t(x)}\tilde{\Gamma}_m^{t(x)}\right](x)$ by definition of the operator Ψ_w. Since this upper bound does not depend on the value of x, it follows that

$$\begin{aligned}\left\|\log \Psi_w[\tilde{\Gamma}_m] - \log \Psi_w[\Gamma_m]\right\|_\infty &= \sup_{x\in\mathcal{X}}\left|\log \Psi_w[\tilde{\Gamma}_m](x) - \log \Psi_w[\Gamma_m](x)\right| \\ &\le \lambda \left\|\log \tilde{\Gamma}_m - \log \Gamma_m\right\|_\infty\end{aligned}$$

and by a similar argument,

$$\left\|\log \Psi_m[\tilde{\Gamma}_w] - \log \Psi_m[\Gamma_w]\right\|_\infty \quad \leq \quad \lambda \left\|\log \tilde{\Gamma}_w - \log \Gamma_w\right\|_\infty$$

Since by Assumption 2.1 and the expression in equation (A.3), $\lambda = \frac{\exp\{\bar{U}+\bar{V}\}}{1+\exp\{\bar{U}+\bar{V}\}} < 1$, the mapping $(\log \Gamma_w, \log \Gamma_m) \mapsto (\log \Psi_w[\Gamma_m], \log \Psi_m[\Gamma_w])$ is indeed a contraction.

Existence and uniqueness of fixed point: Since we showed in the first step that the solution (Γ_w^*, Γ_m^*), if it exists, has to be continuous, we can take the fixed point mapping $(\log \Gamma_w, \log \Gamma_m) \mapsto (\log \Psi_w[\Gamma_m], \log \Psi_m[\Gamma_w])$ to be its restriction to the space of continuous functions $(\mathcal{C}^* \times \mathcal{C}^*, \|\cdot\|_\infty)$ endowed with the supremum norm

$$\|(\Gamma_w, \Gamma_m)\|_\infty := \max\left\{\sup_x |\log \Gamma_w(x)|, \sup_z |\log \Gamma_m(z)|\right\}.$$

Since this space is a complete vector space, and $(\log \Psi_w, \log \Psi_m)$ is a contraction mapping, the conclusion follows directly using Banach's fixed point theorem. □

Appendix B. Proofs for Section 4

In the following, let $\tilde{U}_{ij} := U(x_i, z_j)$ and $\tilde{U}_{ik} := U(x_i, z_k)$. Before proving Theorem 4.1, we are going to establish the following Lemma:

Lemma B.1. *Suppose that Assumptions 2.1, 2.2, and 4.1 hold, and that the random utilities* $U_{i0}, U_{i1}, \dots, U_{iJ}$ *are* J *i.i.d. draws from the model in (2.1) and (2.2). Then as* $J \to \infty$,

$$\left|P\left(U_{i0} \geq U_{ik}, k = 0, \dots, J | \tilde{U}_{i1}, \dots \tilde{U}_{iJ}\right) - \frac{1}{1 + \frac{1}{J}\sum_{k=1}^J \exp\{\tilde{U}_{ik}\}}\right| \quad \to \quad 0, \quad \text{and}$$

$$\left|JP\left(U_{ij} \geq U_{ik}, k = 0, \dots, J | \tilde{U}_{i1}, \dots \tilde{U}_{iJ}\right) - \frac{\exp\{\tilde{U}_{ij}\}}{1 + \frac{1}{J}\sum_{k=1}^J \exp\{\tilde{U}_{ik}\}}\right| \quad \to \quad 0$$

for any fixed $j = 1, 2, \dots, J$.

Proof: Using independence, the conditional probability that $U_{ij} \geq U_{ik}$ for all $k = 1, \dots, J$ given η_{ij} is equal to

$$P\left(U_{ij} \geq U_{ik}, k = 1, \dots, J | \tilde{U}_{i1}, \dots \tilde{U}_{iJ}, \eta_{ij}\right) = \prod_{k \neq j} G(\eta_{ij} + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik}))$$

for any $j = 0, 1, \dots, J$. By the law of iterated expectations, we the unconditional probability is obtained by integrating over the density of η,

$$\begin{aligned} P\left(U_{ij} \geq U_{ik}, k = 1, \dots, J | \tilde{U}_{i1}, \dots \tilde{U}_{iJ}\right) &= \int_{-\infty}^{\infty} \left[\prod_{k \neq j} G(s + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik}))\right] g(s) ds \\ &= \int_{-\infty}^{\infty} \exp\left\{\sum_{k \neq j} \log G(s + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik}))\right\} g(s) ds \\ &= \int_{-\infty}^{\infty} \exp\left\{\sum_{k=0}^J \log G(s + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik}))\right\} \frac{g(s)}{G(s)} ds \quad \text{(B.1)} \end{aligned}$$

where the last step follows since $\tilde{U}_{ij} - \tilde{U}_{ij} = 0$. Now we can rewrite the exponent in the last expression as

$$\sum_{k=0}^{J} \log G(s + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik})) = \frac{1}{J}\sum_{k=0}^{J} J \log \left(G(s + \sigma_J^{-1}(\tilde{U}_{ij} - \tilde{U}_{ik}))\right)$$

We now let the sequences $b_J := G^{-1}\left(1 - \frac{1}{J}\right) \to \infty$ and $a_J = a(b_J) = \sigma_J^{-1}$ - where $a(z)$ is the auxiliary function specified in Assumption 2.2. Then, by a change of variables $s = a_J t + b_J$, we can rewrite the integral in (B.1) as

$$P\left(U_{ij} \geq U_{ik}, k = 1, \dots, J | \tilde{U}_{i1}, \dots \tilde{U}_{iJ}\right) = \int_{-\infty}^{\infty} \exp\left\{\frac{1}{J}\sum_{k=0}^{J} J \log G(a_J(t + \tilde{U}_{ij} - \tilde{U}_{ik}) + b_J)\right\} \frac{a_J g(a_J t + b_J)}{G(a_J t + b_J)} dt$$

Convergence of the Integrand. We next show that for $j \neq 0$, the integrand converges to a non-degenerate limit as $J \to \infty$. First consider the exponent

$$R_J(t) := \frac{1}{J}\sum_{k=0}^{J} J \log G(a_J(t + \tilde{U}_{ij} - \tilde{U}_{ik}) + b_J)$$

Since for $G \to 1$, $-\log G \approx 1 - G$, we obtain

$$\begin{aligned} R_J(t) &= -\frac{1}{J}\sum_{k=0}^{J} J(1 - G(b_J + a_J(t + \tilde{U}_{ij} - \tilde{U}_{ik}))) + o(1) \\ &= -\frac{1}{J}\sum_{k=0}^{J} J(1 - G(b_J + a(b_J)(t + \tilde{U}_{ij} - \tilde{U}_{ik}))) + o(1) \end{aligned}$$

where the last step follows from the choice of a_J. Since $(1 - G(s))^{-1}$ is Γ-varying with auxiliary function $a(s)$, and $b_J \to \infty$,

$$\frac{1 - G(b_J + a(b_J)(t + \tilde{U}_{ij} - \tilde{U}_{ik}))}{1 - G(b_J)} \to \exp\{-t - (\tilde{U}_{ij} - \tilde{U}_{ik})\}$$

Finally, since $G(b_J) = 1 - \frac{1}{J}$,

$$J(1 - G(b_J + a(b_J)(t + \tilde{U}_{ij} - \tilde{U}_{ik}))) = \frac{(1 - G(b_J + a(b_J)(t + \tilde{U}_{ij} - \tilde{U}_{ik})))}{1 - G(b_J)} \to \exp\{-t - (\tilde{U}_{ij} - \tilde{U}_{ik})\}$$

Since $G(a_J t + b_J)$ is also nondecreasing in t, convergence of the integrand is also locally uniform with respect to t and $(\tilde{U}_{ij} - \tilde{U}_{ik})$ by the arguments in section 0.1 in Resnick (1987). Hence,

$$R_J(t) = -e^{-t}\frac{1}{J}\sum_{k=0}^{J} \exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\} + o(1) \tag{B.2}$$

where the term $\frac{1}{J}\sum_{k=0}^{J} \exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\} \leq \exp\{2\bar{U}\} < \infty$ is uniformly bounded by Assumption 2.1. Next, we turn to the term

$$r_J(t) = J a_J g(b_J + a_J t)$$

Since $a_J = a(b_J)$, and $a(z) = \frac{1-G(z)}{g(z)}$, we can write

$$r_J(t) = J a(b_J) \frac{1 - G(b_J a_J t)}{a(b_J + a_J t)}$$

By the same steps as before,

$$J(1 - G(b_J + a_J t)) \to e^{-t}$$

Furthermore by Lemma 1.3 in Resnick (1987), we have

$$\frac{a(b_J)}{a(b_J + a_J t)} \to 1$$

so that

$$r_J(t) \to e^{-t}$$

Combining this result with (B.2), we get

$$\begin{aligned} J \quad & \exp\left\{\frac{1}{J}\sum_{k=0}^{J} J \log G(a_J(t + \tilde{U}_{ij} - \tilde{U}_{ik}) + b_J)\right\} a_J g(a_J t + b_J) = \exp\{R_J(t)\} r_J(t) \\ \to \quad & \exp\left\{-t - e^{-t}\frac{1}{J}\sum_{k=0}^{J}\exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\}\right\} \end{aligned}$$

for every $t \in \mathbb{R}$.

Convergence of the Integral. Let $h_J^*(t) := \exp\left\{-t - e^{-t}\frac{1}{J}\sum_{k=0}^{J}\exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\}\right\}$. Since the function $h_J(t) := \exp\{R_J(t)\} r_J(t)$ is bounded uniformly in J, and $|h_J(t) - h_J^*(t)| \to 0$ pointwise, it follows that

$$\left| JP\left(U_{ij,n} \geq U_{ik,n}, k = 0, \ldots, J | \tilde{U}_{i1,n}, \ldots \tilde{U}_{iJ,n}\right) - \int_{-\infty}^{\infty} h_J^*(t)dt\right| = \left|\int_{-\infty}^{\infty}(h_J(t) - h_J^*(t))dt\right| \to 0$$

from the Dominated Convergence Theorem. From a change in variables $\psi := -e^{-t}$, we can evaluate the integral

$$\begin{aligned} \int_{-\infty}^{\infty} h_J^*(t)dt &= \int_{-\infty}^{\infty}\exp\left\{-e^{-t}\frac{1}{J}\sum_{k=0}^{J}\exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\}\right\} e^{-t}dt \\ &= \left(\frac{1}{J}\sum_{k=0}^{J}\exp\{\tilde{U}_{ik} - \tilde{U}_{ij}\}\right)^{-1} = \frac{\exp\{\tilde{U}_{ij}\}}{\frac{1}{J}\sum_{k=0}^{J}\exp\{\tilde{U}_{ik}\}} \\ &= \frac{\exp\{\tilde{U}_{ij}\}}{1 + \frac{1}{J}\sum_{k=1}^{J}\exp\{\tilde{U}_{ik}\}} \end{aligned}$$

since $U_{i0}^* = \log J$. Hence,

$$\left| JP\left(U_{ij} \geq U_{ik}, k = 0, \ldots, J | \tilde{U}_{i1}, \ldots \tilde{U}_{iJ}\right) - \frac{\exp\{\tilde{U}_{ij}\}}{1 + \frac{1}{J}\sum_{j=1}^{J}\exp\{\tilde{U}_{ik}\}}\right| \to 0$$

for each $j = 1, 2, \ldots, J$, as claimed in (B.1). Furthermore, it follows that

$$\begin{aligned} P\left(U_{i0} \geq U_{ik}, k = 0, \ldots, J | \tilde{U}_{i1}, \ldots \tilde{U}_{iJ}\right) &= 1 - \sum_{j=1}^{J} P(U_{ij} \geq U_{ik}, k = 0, \ldots, J) \\ &= \frac{1}{1 + \frac{1}{J}\sum_{j=1}^{J}\exp\{\tilde{U}_{ik}\}} + o(1) \end{aligned} \tag{B.3}$$

which establishes the second assertion. □

B.1. **Proof of Theorem 4.1.** For the main conclusion of the theorem, note that since $z_1, z_2, \ldots$ are a sequence of i.i.d. draws from $M(z)$, Assumption 2.1 and a law of large numbers can be used to establish

$\frac{1}{J}\sum_{k=1}^{J}\exp\{U(x_i,z_j)\} \to \int \exp\{U(x_i,s)\}m(s)ds$. It follows by the continuous mapping theorem that

$$\frac{\exp\{U(x_i,z_j)\}}{1+\frac{1}{J}\sum_{k=1}^{J}\exp\{U(x_i,z_k)\}} \to \frac{\exp\{U(x_i,z_j)\}}{1+\int \exp\{U(x_i,s)\}m(s)ds}$$

almost surely, so that the conclusion of Theorem 4.1 follows from Lemma B.1 and the triangle inequality. □

B.2. **Auxiliary Lemmas for the Proof of Theorem 4.2.** To prove the main result in Theorem 4.2, we are first going to establish two technical Lemmas. In calculating conditional distributions, we have to account for the multiplicity of stable matchings. We will do this by showing that bounds on the relevant joint and conditional c.d.f.s - which do not depend on how a matching is selected from the set of stable matchings - shrink towards a singleton at a sufficiently fast rate.

The first result concerns the rate at which the number of available potential spouses increases for each individual in the market. For a given stable matching μ^*, we let

$$J_{wi}^* = \sum_{j=1}^{n} \mathbb{1}\left\{V_{ji} \geq V_j^*(W_j^*)\right\} \qquad \text{and } J_{mj}^* = \sum_{i=1}^{n} \mathbb{1}\left\{U_{ij} \geq U_i^*(M_i^*)\right\}$$

denote the number of men available to woman i, and the number of women available to man j, respectively, where M_i^* and W_j^* denote woman i's and man j's opportunity sets under μ^*, and $U_i^*(M) := \max_{j\in M} U_{ij}$ and $V_j^*(W) := \max_{i\in W} V_{ji}$, where by convention, the outside option $0 \in W_j^*$ and $0 \in M_i^*$. Similarly, we let

$$L_{wi}^* = \sum_{j=1}^{n} \mathbb{1}\left\{U_{ij} \geq U_i^*(M_i^*)\right\} \qquad \text{and } L_{mj}^* = \sum_{i=1}^{n} \mathbb{1}\left\{V_{ji} \geq V_j^*(W_j^*)\right\}$$

so that L_{wi} is the number of men to whom woman i is available, and L_{mj} the number of women to whom man j is available. Lemma B.2 below establishes that in our setup, the number of available potential matches grows at a root-n rate as the size of the market grows.

Lemma B.2. *Suppose Assumptions 2.1, 2.2, and 4.1 hold, then under any stable matching, (a)*

$$\begin{aligned} n^{1/2}\frac{\exp\{-\bar V+\gamma_m\}}{1+\exp\{\bar U+\bar V+\gamma_w\}} \leq &\quad J_{wi}^* &\leq n^{1/2}\exp\{\bar V+\gamma_m\} \\ n^{1/2}\frac{\exp\{-\bar U+\gamma_w\}}{1+\exp\{\bar U+\bar V+\gamma_m\}} \leq &\quad J_{mj}^* &\leq n^{1/2}\exp\{\bar U+\gamma_w\} \end{aligned}$$

for each $i=1,\dots,n$ and $j=1,\dots,n$ with probability approaching 1 as $n\to\infty$. (b) Furthermore,

$$\begin{aligned} n^{1/2}\frac{\exp\{-\bar U+\gamma_m\}}{1+\exp\{\bar U+\bar V+\gamma_m\}} \leq &\quad L_{wi}^* &\leq n^{1/2}\exp\{\bar U+\gamma_m\} \\ n^{1/2}\frac{\exp\{-\bar V+\gamma_w\}}{1+\exp\{\bar U+\bar V+\gamma_w\}} \leq &\quad L_{mj}^* &\leq n^{1/2}\exp\{\bar V+\gamma_w\} \end{aligned}$$

for each $i=1,\dots,n$ and $j=1,\dots,n$ with probability approaching 1 as $n\to\infty$.

Proof: First note that by Assumption 2.1 $U(x,z)$ and $V(z,x)$ are bounded, we can bound choice probabilities of the form $P(U_{ij}\leq b_n)$ by

$$P(\eta_{ij}\leq b_n-2\bar U)\leq P(U_{ij}\leq b_n)\leq P(\eta_{ij}\leq b_n+2\bar U)$$

where $\bar U<\infty$.

Since the sets of available spouses W_i^* and M_j^* under the stable matching are endogenous, the taste shifters η_{ij} and ζ_{ji} are in general not independent conditional on those choice sets. To circumvent this difficulty, the following argument only relies on lower and upper bounds on U_i^* and V_i^* that are implied by the respective utilities of the outside option U_{i0}, V_{j0}, and unconditional independence of taste shocks.

Rate for Expectation of upper bound $\bar{J}_{mj}$. In the following, we denote the set of women that prefers man j to their outside option by $\bar{W}_j$. Since every woman can choose to remain single, we can bound J_{mj}^* by

$$\begin{aligned} J_{mj}^* &= \sum_{i=1}^{n_w} \mathbb{1}\left\{i \in W_j^*\right\} = \sum_{i=1}^{n_w} \mathbb{1}\left\{U_{ij} \geq U_i^*(M_i^*)\right\} \\ &\leq \sum_{i=1}^{n_w} \mathbb{1}\left\{U_{ij} \geq U_{i0}\right\} = \sum_{i=1}^{n_w} \mathbb{1}\left\{i \in \bar{W}_j\right\} =: \bar{J}_{mj} \end{aligned}$$

By Assumption 4.1 and Lemma B.1,

$$JP\left(U_{ij} \geq U_{i0} | x_i, z_j\right) \to \frac{\exp\{\tilde{U}_{ij}\}}{1 + \frac{1}{J}\exp\{\tilde{U}_{ij}\}}$$

Hence, we can obtain the expectation of the upper bound J_{mj}^*,

$$\mathbb{E}[\bar{J}_{mj}|z_j, x_1, \dots, x_n] \to \frac{1}{J}\sum_{i=1}^{n_w} \frac{\exp\{\tilde{U}_{ij}\}}{1 + \frac{1}{J}\exp\{U_{ij}\}} \leq \frac{n_w}{J}\exp\{\bar{U}\}$$

where $\bar{U} < \infty$ was given in Assumption 2.1. Since by Assumption 4.1, $J = [n^{1/2}]$ and the bound on the right-hand side does not depend on $z_j, x_1, \dots, x_n$, we have, by the law of iterated expectations that

$$\mathbb{E}[\bar{J}_{mj}] \leq n^{1/2}(\exp\{\bar{U}\} + \gamma_w + o(1)) \tag{B.4}$$

where the remainder term $o(1)$ can be shown to converge uniformly for $j = 1, 2, \dots$.

Rate for Variance $\bar{J}_{mj}$. Let $p_{ijn} := \frac{\exp\{\tilde{U}_{ij}\}}{J+\exp\{\tilde{U}_{ij}\}}$ and $\bar{v}_{jn} := \frac{1}{n}\sum_{i=1}^{n_w} p_{ijn}(1 - p_{ijn})$. Since by Assumption 4.1, $p_{ijn} \leq \frac{\exp\{\bar{U}\}}{\sqrt{n}+1+\exp\{\bar{U}\}}$ and $p_{ijn} \geq \frac{\exp\{-\bar{U}\}}{\sqrt{n}+1+\exp\{-\bar{U}\}}$, we have that $(n^{1/2} + 2)^{-1}\exp\{-\bar{U} + \gamma_w\} \leq \bar{v}_{jn} \leq n^{-1/2}\exp\{\bar{U} + \gamma_w\}$. Hence, $\bar{v}_{jn} \to 0$ and $n\bar{v}_{jn} \to \infty$.

Since $\eta_{i0,k}$, $k = 1, \dots, J$ are i.i.d. draws from the distribution $G(\eta)$, we can apply a CLT for independent heterogeneously distributed random variables to the upper bound $\bar{J}_{mj}$,

$$\frac{\bar{J}_{mj} - \mathbb{E}[\bar{J}_{mj}]}{\sqrt{n\bar{v}_{jn}}} = \frac{1}{\sqrt{n\bar{v}_{jn}}}\sum_{i=1}^{n_w}\left(\mathbb{1}\{U_{ij} \geq U_{i0}\} - p_{ijn}\right) \stackrel{d}{\to} N(0, 1)$$

where the Lindeberg condition holds since the random variables $\mathbb{1}\{U_{ij} \geq U_{i0}\}$ are bounded, and $n\bar{v}_{jn} \to \infty$. Since $\bar{v}_{jn} \to 0$ uniformly in $j = 1, 2, \dots$, we obtain that

$$\frac{\bar{J}_{mj} - \mathbb{E}[\bar{J}_{mj}]}{\sqrt{n}} \stackrel{p}{\to} 0$$

uniformly in $j = 1, 2, \dots$.

Rate for Expectation of lower bound J_{wi}°. Next, we denote the set of men j that prefer woman i to their outside option or any woman in $\bar{W}_j$ by M_i°. Since by construction, $\bar{W}_j$ is a superset of (i.e. contains) W_j^*, $M_i^\circ \subset M_i^*$. Hence, we can bound J_{wi}^* by

$$\begin{aligned} J_{wi}^* &= \sum_{j=1}^{n_m} \mathbb{1}\left\{j \in M_i^*\right\} = \sum_{j=1}^{n_m} \mathbb{1}\left\{V_{ji} \geq V_j^*(W_j^*)\right\} \\ &\geq \sum_{j=1}^{n_m} \mathbb{1}\left\{V_{ji} \geq V_j^*(\bar{W}_j)\right\} = \sum_{j=1}^{n_m} \mathbb{1}\left\{j \in M_i^\circ\right\} =: J_{wi}^\circ \end{aligned}$$

Applying Lemma B.1 again, we obtain

$$JP\left(V_{ji} \geq \max_{k \in \bar{W}_j} V_{jk} \middle| x_i, z_j, \bar{W}_j\right) \to \frac{\exp\{\tilde{V}_{ji}\}}{1 + \frac{1}{J}\sum_{k \in \bar{W}_j} \exp\{\tilde{V}_{jk}\}} \geq \frac{J\exp\{-\bar{V}\}}{J + \bar{J}_{mj}\exp\{\bar{V}\}}$$

where $\bar{V} < \infty$ was defined in Assumption 2.1.

Finally, note that this lower bound is a convex function of $\bar{J}_{mj}$, so that we can use our previous bound in (B.5) together with Jensen's Inequality to obtain

$$JP\left(V_{ji} \geq \max_{k \in \bar{W}_j} V_{jk} \middle| x_i, z_j\right) \geq \frac{J\exp\{-\bar{V}\}}{J + \mathbb{E}[\bar{J}_{mj}]\exp\{\bar{V}\}} \geq \frac{J\exp\{-\bar{V}\}}{J + n^{1/2}\exp\{\bar{U} + \bar{V}\}}$$

which is bounded for all values of J since $J = [\sqrt{n}]$. Hence, by the law of iterated expectations, we can obtain the expectation of the lower bound J°_{wi},

$$\mathbb{E}[J^\circ_{wi}] = \sum_{j=1}^{n_m} P\left(V_{ji} \geq \max_{k \in \bar{W}_j} V_{jk} | x_i, z_j\right) \geq n^{1/2}\left(\frac{\exp\{-\bar{V} + \gamma_m\}}{1 + \exp\{\bar{U} + \bar{V} + \gamma_w\}} + o(1)\right) \quad \text{(B.5)}$$

where the remainder term $o(1)$ can be shown to converge uniformly for $i = 1, 2, \ldots$.

Rate for Variance J°_{wi}. Let $p_{jin} := \frac{\exp\{\tilde{V}_{ij}\}}{J + \sum_{k \in \bar{W}_j} \exp\{\tilde{V}_{ik}\}}$ and $\bar{v}_{in} := \frac{1}{n}\sum_{j=1}^{n_m} p_{jin}(1 - p_{jin})$. Using the corresponding bounds derived above and similar steps as for $\bar{v}_{jn}$, we obtain $\bar{v}_{in} \to 0$ and $n\bar{v}_{in} \to \infty$. Since $\zeta_{j0,k}$, $k = 1, \ldots, J$, and ζ_{ji}, $i = 1, \ldots, n$ are i.i.d. draws from the distribution $G(\eta)$ and independent of $\bar{W}_j$, we can again apply the Lindeberg-Lévy CLT to obtain that

$$\frac{J^\circ_{wi} - \mathbb{E}[J^\circ_{wi}]}{\sqrt{n}} \xrightarrow{p} 0$$

uniformly in $i = 1, 2, \ldots$.

Symmetry: Bounds for both sides. If we reverse the role of the male and female sides of the market, we can repeat the same sequence of steps and obtain a lower bound $J^\circ_{mj} \leq J^*_{mj}$ and an upper bound $\bar{J}_{wi} \geq J^*_{wi}$ satisfying

$$\begin{aligned} \mathbb{E}[\bar{J}_{wi}] &\leq n^{1/2}(\exp\{\bar{V} + \gamma_m\} + o(1)) \\ \mathbb{E}[J^\circ_{mj}] &\geq n^{1/2}\left(\frac{\exp\{-\bar{U} + \gamma_w\}}{1 + \exp\{\bar{U} + \bar{V} + \gamma_m\}} + o(1)\right) \end{aligned}$$

where

$$\frac{\bar{J}_{wi} - \mathbb{E}[\bar{J}_{wi}]}{\sqrt{n}} \xrightarrow{p} 0 \quad \text{and} \quad \frac{J^\circ_{mj} - \mathbb{E}[J^\circ_{mj}]}{\sqrt{n}} \xrightarrow{p} 0$$

which concludes the proof of part (a). The proof of part (b) is completely analogous. □

Next, we show that the dependence between η_{ij}, ζ_{ji}, and opportunity sets becomes small as n increases. In the following we use indices i and k to denote a specific (generic, respectively) woman in the market, and j and l to denote a specific (generic) man. The indicator variable D_{il} is equal to 1 if man l is available to woman i, and zero otherwise. Similarly, we let E_{jk} be an indicator variable that is equal to 1 if woman k is available to j, and zero otherwise. We are going to consider the joint distribution of $\eta_i := (\eta_{i1}, \ldots, \eta_{in_m})'$, $\zeta_j = (\zeta_{ji}, \ldots, \zeta_{jn_w})'$ and the indicator variables $D^*_{il} := \mathbb{1}\{l \in M^*_i\}$, and $E^*_{jk} = \mathbb{1}\left\{k \in W^*_j\right\}$ for different stable matchings. We also let $D^*_{i,-j} := (D^*_{i1}, \ldots, D^*_{i(j-1)}, D^*_{i(j+1)}, \ldots, D^*_{in_m})$ and $E^*_{j,-i} := (E^*_{j1}, \ldots, E^*_{j(i-1)}, E^*_{j(i+1)}, \ldots, E^*_{jn_w})$ for any stable matching, and use analogous notation for the M- and W-preferred matchings. Then for any vectors of indicator variables $\mathbf{d} = (d_1, \ldots, d_{n_m-1}) \in \{0,1\}^{n_m-1}$

and $\mathbf{e} = (e_1, \ldots, e_{n_w-1}) \in \{0,1\}^{n_w-1}$ we denote the conditional c.d.f.s

$$\begin{aligned} G^*_{\eta|D^*}(\eta|\mathbf{d}) &= P\left(\eta_i \leq \eta | D^*_i = \mathbf{d} \text{ for some stable matching}\right) \\ G^*_{\eta,\zeta|D^*,E^*}(\eta,\zeta|\mathbf{d},\mathbf{e}) &= P\left(\eta_i \leq \eta, \zeta_j \leq \zeta | D^*_{i,-j} = \mathbf{d}, E^*_{j,-i} = \mathbf{e} \text{ for some stable matching}\right) \end{aligned}$$

with the associated p.d.f.s $g^*_{\eta|D^*}(\eta,\zeta|\mathbf{d})$ and $g^*_{\eta,\zeta|D^*,E^*}(\eta,\zeta|\mathbf{d},\mathbf{e})$, respectively. In words, $G^*(\eta,\zeta|\mathbf{d},\mathbf{e})$ denotes the conditional distribution of taste shifters η_i and ζ_j given that the opportunity sets M_i and W_j corresponding to the indicator variables $\mathbf{d},\mathbf{e}$ is supported by at least one stable matching.

We can now state the following lemma characterizing the conditional distribution of taste shifters given an agent's opportunity set.

Lemma B.3. *Suppose Assumptions 2.1, 2.2, and 4.1 hold. Then (a) the conditional distribution for η given that D^*_i is supported by a stable matching satisfies*

$$\lim_n \left| \frac{g^*_{\eta|D^*}(\eta|D^*_i)}{g_\eta(\eta)} - 1 \right| = 0$$

with probability approaching one as $n \to \infty$. The analogous results hold for the male side of the market. Furthermore, (b) the conditional distributions for (η,ζ) given $D^*_{i,-j}, E^*_{j,-i}$ satisfies

$$\lim_n \left| \frac{g^*_{\eta,\zeta|D^*,E^*}(\eta,\zeta|D^*_{i,-j},E^*_{j,-i})}{g_{\eta,\zeta}(\eta,\zeta)} - 1 \right| = 0$$

with probability approaching one as $n \to \infty$. (c) The analogous conclusion holds for any finite subset of men $M_0 \subset \{1,\ldots,n_m\}$ and women $W_0 \subset \{1,\ldots,n_w\}$, where the conditioning set excludes the availability indicators between any pair $k \in W_0$ and $l \in M_0$.

Proof: Without loss of generality, let $\gamma_w = \gamma_m = 0$. We first prove part (a). Let B denote the event that the opportunity set $d_1,\ldots,d_{n_m}$ for woman i is supported by a stable matching. Given the systematic parts of payoffs, $U(x_k,z_l)$ and $V(z_l,x_k)$ for women $k=1,\ldots,n_w$ and men $l=1,\ldots,n_m$, this event is determined by the values of the idiosyncratic taste shifters $\eta_1,\ldots,\eta_{n_w}$ and $\zeta_1,\ldots,\zeta_{n_m}$.

We need to establish that the conditional distribution of η_i given B converges to the unconditional distribution at a sufficiently fast rate. To that end, we consider the probability densities $g^*_{\eta|B}(\eta_i|B)$ and $g^*_{\tilde\eta|B}(\tilde\eta_i|B) = g_\eta(\tilde\eta_i)$. By the definition of conditional densities, we can write

$$\frac{g^*_{\tilde\eta|B}(\eta|B)}{g^*_{\eta|B}(\eta|B)} = \frac{g^*_{\tilde\eta,B}(\eta,B)P(B)}{g^*_{\eta,B}(\eta,B)P(B)} = \frac{P(B|\tilde\eta=\eta)g_{\tilde\eta}(\eta)}{P(B|\eta_i=\eta)g_\eta(\eta)} = \frac{P(B|\tilde\eta_i=\eta)}{P(B|\eta_i=\eta)} \tag{B.6}$$

where the last step follows since the marginal distributions of $g_\eta(\eta_i)$ and $g_{\tilde\eta}(\tilde\eta_i)$ are the same by construction.

We then find a common bound on the relative change in the conditional probability of B given η_i that is independent of B and η_i, and apply that bound to the event B to establish that $\left|\frac{P(B|\eta_i)}{P(B|\tilde\eta_i)} - 1\right| \to 0$ almost surely. Hence, as a final step it follows from (B.6) that $\left|\frac{g^*_{\eta|B}(\eta|B)}{g^*_{\tilde\eta}(\eta|B)} - 1\right| \to 0$ so that

$$\left| \frac{g^*_{\eta|B}(\eta_i|B)}{g_\eta(\eta_i)} - 1 \right| = \left| \frac{g^*_{\eta|B}(\eta_i|B)}{g^*_{\tilde\eta|B}(\tilde\eta_i|B)} - 1 \right| \to 0$$

where the equality follows from the fact that by construction $\tilde\eta_{i1},\ldots,\tilde\eta_{in_m}$ are jointly independent of B and have the same marginal distribution as $\eta_{i1},\ldots,\eta_{in_m}$.

Now let $\bar{\eta} = (\eta_1', \dots, \eta_{n_w}')'$, $\bar{\zeta} = (\zeta_1', \dots, \zeta_{n_m}')'$, $\bar{\eta}_{-i} = (\eta_1', \dots, \eta_{i-1}', \eta_{i+1}', \dots, \eta_{n_w}')'$, and define the random variable

$$I(\eta, B) := \mathbb{1}\left\{\bar{\eta}_{-i}, \eta_i = \eta, \bar{\zeta} \text{ support } B\right\}$$

an indicator whether B results from a stable matching given the realizations of taste shifters. We can then write

$$P(B|\eta_i) = \int I(\eta_i, B) dG(\bar{\eta}, \bar{\zeta}|\eta_i) = \int I(\eta_i, B) dG(\bar{\eta}, \bar{\zeta})$$

since η_i and $\bar{\eta}_{-i}, \bar{\zeta}$ are (unconditionally) independent by assumption.

The remainder of this proof establishes that

$$\begin{aligned}
\frac{P(B|\eta_i) - P(B|\tilde{\eta}_i)}{P(B|\eta_i)} &= \frac{\int (I(\eta_i, B) - I(\tilde{\eta}_i, B)) dG(\bar{\eta}_{-i}, \bar{\zeta})}{P(B|\eta_i)} \\
&\leq \frac{\int I(\eta_i, B)(1 - I(\tilde{\eta}_i, B)) dG(\bar{\eta}_{-i}, \bar{\zeta})}{P(B|\eta_i)} \\
&= \int (1 - I(\tilde{\eta}_i, B)) dG(\bar{\eta}_{-i}, \bar{\zeta}|B, \eta_i)
\end{aligned}$$

and

$$\frac{\int (I(\tilde{\eta}_i, B) - I(\eta_i, B)) dG(\bar{\eta}_{-i}, \bar{\zeta})}{P(B|\tilde{\eta}_i)} \leq \int (1 - I(\eta_i, B)) dG(\bar{\eta}_{-i}, \bar{\zeta}|B, \tilde{\eta}_i)$$

both converge to zero as n_m and n_w grow large. Combining these two statements, it then follows that $\left|\frac{P(B|\eta_i)}{P(B|\tilde{\eta}_i)} - 1\right| \to 0$.

Conditional and Unconditional Probability of B. Next, we iteratively bound the probability that replacing the taste shifters η_i with an independent copy $\tilde{\eta}_i$ alters woman i's opportunity set. We use a tilde to distinguish the value of a variable resulting from such a change from its original value. Specifically, we let $\tilde{D}_{kl}^{(s)}$ denote the indicator whether man l is available to woman k after the sth iteration, $\tilde{E}_{lk}^{(s)}$ an indicator whether woman k is available to man l after the sth iteration, and $\tilde{U}_k^{*(s)}$ and $\tilde{V}_l^{*(s)}$ woman k and man l's respective indirect utility given their opportunity sets at the sth stage. We also define the event $\tilde{B} := \{\tilde{D}_{i1}^{(\infty)} = d_{i1}, \dots, \tilde{D}_{in_m}^{(\infty)} = d_{in_m}\}$.

Next, we show that replacing all of i's taste shifters in an arbitrary fashion starts two parallel "chains" of subsequent changes, where at each iteration, there is at most one element in each of the two sets of dummies $\left\{\tilde{D}_{kl}^{(s)}\right\} k,l$ and $\left\{\tilde{E}_{kl}^{(s)}\right\}_{k,l}$ that will be changed at the sth stage and has an impact on subsequent rounds. Furthermore, at each iteration, there is a nontrivial probability that the shift in the previous iteration only affects the outside option, in which case the chain will be terminated at that stage.

If there are no further adjustments after stage s, the resulting indicators $\tilde{D}_{kl}^{(s)}$ and $\tilde{E}_{lk}^{(s)}$ satisfy the pairwise stability conditions since the initial opportunity sets D_{kl} and E_{lk} correspond to a pairwise stable matching given the taste shifters $\bar{\eta}, \bar{\zeta}$ by assumption. Hence the resulting opportunity set for woman i $\tilde{D}_{i1}^{(s)}, \dots, \tilde{D}_{in_m}^{(s)}$ is supported by a stable matching given the new taste shifters $\tilde{\eta}_i, \bar{\eta}_{-i}, \bar{\zeta}$.

For the following arguments, note that the event B only contains information regarding the number and identities of men with $D_{il} = 1$. Hence, the conditional distribution of $\bar{\eta}, \bar{\zeta}$ given B is invariant with respect to permutations of identities among (a) the set of women other than i, (b) the set of men with $D_{il} = 1$, and (c) the set of men with $D_{il} = 0$, respectively. Furthermore, since the new taste shifters $\tilde{\eta}_{ij}$ were drawn independently of B, the identity of the "switchers" among the men with $D_{ij} = 1$ is also independent of B conditional on characteristics x_k, z_l.

Base Case. We now consider the direct effect of replacing η_i with $\tilde{\eta}_i$ on whether i is available to j, holding all other taste shifters fixed. Replacing η_i with an independent copy $\tilde{\eta}_i$ changes U_i^* to $\tilde{U}_i^{*(1)} :=$

$\max\{U(x_i,z_j)+\sigma_n\tilde{\eta}_{ij}: D_{ij}=1\}$. Next note that if $D_{ij}=0$, then $V_{ji}<V_j^*$, so that any changes to E_{ij} do not have any subsequent effects on j's choices and can therefore be ignored. On the other hand, if $D_{ij}=1$, then $V_{ji}\geq V_j^*$, so that a change from $E_{ji}=0$ to $\tilde{E}_{ji}^{(1)}=1$ increases j's indirect utility. Hence if for woman k, $V_{ji}>V_{jk}>V_j^*$, we have that $D_{kj}=1$ and $\tilde{D}_{kj}^{(2)}=0$. Hence it is sufficient to consider shifts in E_{ji} for men j such that $D_{ij}=1$.

Among the subset of men l such that $D_{il}=1$, with probability 1 there is exactly one man j such that $E_{ji}=1$ and exactly one man j' such that $\tilde{E}_{j'i}^{(1)}=1$, where j and j' can also be the indices corresponding to the outside option. Hence, there are at most two switches of the form $E_{ji}\neq\tilde{E}_{ji}^{(1)}$ among the subset of men who prefer i to their next best alternative. If j or j' do not correspond to the outside option, we have $D_{ij}=D_{ij'}=1$, so that $\tilde{V}_j^{*(1)}\leq V_{ij}=V_j^*$, and $\tilde{V}_{j'}^{*(1)}=V_{ij'}\geq V_{j'}^*$, respectively.

Inductive Step. We now use induction to show that there is at most one such adjustment at each subsequent round $s=2,3,\dots$: Suppose that after s iterations of one of the two chains, the availability indicators are given by $\tilde{D}_{kl}^{(s)}$ and $\tilde{E}_{lk}^{(s)}$, where $k=1,\dots,n_w$, and $l=1,\dots,n_m$. Under the inductive hypothesis, at the sth stage there was a specific woman i, and among all men $l=1,\dots,n_m$ which were available to a specific woman i (i.e. $\tilde{D}_{il}^{(s)}=1$), there was at most one change of an indicator $\tilde{E}_{ji}^{(s-1)}$ to a new value $\tilde{E}_{ji}^{(s)}$.

Consider first that the last change from $\tilde{E}_{ji}^{(s-1)}=1$ to $\tilde{E}_{ji}^{(s)}=0$. It follows that $\tilde{V}_j^{*(s)}=\max\left\{V_{jk}:\tilde{E}_{jk}^{(s)}=1\right\}=: V_{jk'}$ for some k' such that $\tilde{E}_{jk'}^{(s)}=1$, where k' is unique with probability one. Hence at the $s+1$th iteration, there is a shift from $\tilde{D}_{k'j}^{(s)}=0$ to $\tilde{D}_{k'j}^{(s+1)}=1$, i.e. j becomes available to k'.

Note that j also becomes available to any woman $\tilde{k}$ for whom $V_j^{*(s-1)}>V_{j\tilde{k}}\geq\tilde{V}_j^{*(s)}$. However, by definition of k', any such $\tilde{k}$ would not have been available to j, i.e. $\tilde{E}_{j\tilde{k}}^{(s)}=0$. Hence for $\tilde{k}$, $U_{\tilde{k}j}<\tilde{U}_{\tilde{k}}^{*(s)}$ so that this change has no effect on subsequent iterations. Note that this includes in particular woman i who became unavailable to j at the previous stage.

Next, consider a change from $\tilde{E}_{ji}^{(s-1)}=0$ to $\tilde{E}_{ji}^{(s)}=1$, where $\tilde{D}_{ij}^{(s)}=1$ and j's indirect utility in the previous round was $\tilde{V}_j^{*(s-1)}=: V_{jk'}$ for some k' with $\tilde{E}_{jk'}^{(s-1)}=1$. Since $\tilde{D}_{ij}^{(s)}=1$, it must be true that $V_{ji}\geq\tilde{V}_j^{*(s-1)}$, so that j may become unavailable to woman k', $\tilde{D}_{ij}^{(s+1)}=0$. On the other hand for any k such that $V_{ji}=\tilde{V}_j^{*(s)}>V_{jk}>\tilde{V}_j^{*(s-1)}$, we must have had $\tilde{E}_{jk}^{s}=0$ by definition of $\tilde{V}_j^{*(s-1)}$. Hence with probability 1, the change in the sth round affects at most one woman with $\tilde{E}_{jk}^{(s)}=1$, whereas for women with $\tilde{E}_{ji}^{(s)}=0$ indirect utility does not depend on whether j is available at round $s+1$, so that there is no effect on subsequent iterations.

Hence there is at most one indicator corresponding to a woman k with $\tilde{E}_{jk}^{(s)}=1$ that changes in the sth round. Interchanging the roles of men and women, an analogous argument yields that there is at most one indicator corresponding to a man l with $\tilde{D}_{kl}^{(s)}=1$ that changes in the second part of the sth round, confirming the inductive hypothesis.

Probability of Terminating Events. Each of the two chains of adjustments can terminate at any given stage s if the change in the previous round only affects the outside option, i.e. if $\tilde{V}_l^{*(t)}=V_{l0}$ or $\tilde{U}_k^{*(t)}=U_{k0}$ at $t=s$ or $t=s-1$. On the other hand if the chain results in a change of $D_{i1},\dots,D_{in_m}$ at a given stage, we ignore any subsequent adjustments and treat such a change as the second terminating event. In the following, we bound the conditional probability for each of these two terminating events given B and that the chain has not terminated before the sth stage.

We first derive a lower bound for the probability that the chain is terminated by the outside option at stage s: By the same reasoning as in the proof of Lemma B.2, man l's opportunity set is contained in the set W_l°, where the taste shifters ζ_{lk} are jointly independent of W_l°, and the size of W_l° is bounded from above by $n^{1/2}\exp\{\bar{U}+\gamma_w\}$ with probability approaching 1. Hence, by Lemma B.1, we have

$$n^{-1/2}\exp(\bar{V}) \geq \frac{\exp(V(z_l,x_k))(1+\exp(\bar{V}-\bar{U}+\gamma_w))}{n^{1/2}(1+2\exp(\bar{V}-\bar{U}+\gamma_w))} \geq P(V_{lk} > V_l^*|x_k,z_l) \geq \frac{\exp(V(z_l,x_k))}{n^{1/2}(1+\exp(\bar{U}+\bar{V}+\gamma_w))}$$

for any k,l if n is sufficiently large.

This implies that as n grows large, the (unconditional) share of men remaining single is bounded from below by $\frac{1}{1+\exp(\bar{U}+\bar{V}+\gamma_w)} =: p_s$ with probability approaching 1. By the law of total probability, that bound also holds conditional on i's opportunity set, $(D_{i1}^*,\dots,D_{in_m}^*)$, with probability arbitrarily close to 1 as n grows. Specifically, for the outside option, $P(V_{l0} > V_l^*|D_{ik}^*,x_k,z_l) \geq \frac{1}{1+\exp(\bar{U}+\bar{V}+\gamma_w)}$. Furthermore, by Lemma B.2 part (b), the number of women to whom man l is available is bounded by

$$\underline{L} := n^{1/2}\frac{\exp\{-\bar{V}+\gamma_w\}}{1+\exp\{\bar{U}+\bar{V}+\gamma_w\}} \leq L_{ml} \leq n^{1/2}\exp\{\bar{V}+\gamma_w\} =: \overline{L}$$

with probability approaching 1.

In order to construct a lower bound on the conditional probability that man l is unmatched given $\tilde{D}_{kl}^{(s)}=1$, we can assume the lower bound for L_j^* if j is unmatched, and the upper bound if j is matched. Then, by Bayes law,

$$\begin{aligned} P(V_{l0} > V_l^*|D_{il}^*,\tilde{D}_{lk}^{(s)}=1,x_k,z_l) &\geq \frac{\underline{L}p_s}{\overline{L}(1-p_s)+\underline{L}p_s} \\ &= \frac{n^{-1/2}\underline{L}}{n^{-1/2}\overline{L}\exp(\bar{U}+\bar{V}+\gamma_w)+n^{-1/2}\underline{L}} \end{aligned}$$

which is strictly greater than zero by Assumptions 2.1 and 4.1. Hence, the probability that the shift is not absorbed by the outside option in the sth step is less than or equal to

$$1-P(V_{l0} > V_l^*|D_{ik}^*,\tilde{D}_{lk}^{(s)}=1,x_k,z_l) \leq \frac{\overline{L}\exp(\bar{U}+\bar{V}+\gamma_w)}{\overline{L}\exp(\bar{U}+\bar{V}+\gamma_w)+\underline{L}} =: \lambda$$

where the bound on the right-hand side does not depend on s and is strictly less than one.

Finally, we construct an upper bound for the probability that the chain leads to a change in the availability indicators $D_{i1},\dots,D_{in_m}$ at stage s. To this end, we can follow the same reasoning as for the choice probability for the outside option, where we use the lower bound on the size of the opportunity set from Lemma B.2. Applying Lemma B.1, we then have

$$P(V_{li} > V_l^*|D_{ik}^*,x_k,z_l) \leq n^{-1/2}\frac{\exp(V(z_l,x_k))(1+\exp(\bar{V}-\bar{U}+\gamma_m))}{1+\exp(\bar{V}-\bar{U}+\gamma_m)+\exp(-\bar{U}-\bar{V}+\gamma_w)} \leq n^{-1/2}\exp\{\bar{V}\} =: q_s$$

for n sufficiently large. Hence, the conditional probability that one of the indicators D_{il}, $l=1,\dots,n_m$ is switched given that the process is still active at the sth stage can be bounded by

$$P(\tilde{D}_{ij}^{(s)} \neq D_{ij}|B,\tilde{D}_{lk}^{(s)}=1,z_l,x_k) \leq \frac{n^{-1/2}\exp\{\bar{V}\}\bar{L}}{n^{-1/2}\exp\{\bar{V}\}\bar{L}+\underline{L}} \leq n^{-1/2}\exp\{\bar{V}\}\frac{\bar{L}}{\underline{L}} =: n^{-1/2}\bar{q}$$

where $\bar{q} < \infty$. Clearly, this upper bound becomes arbitrarily small as n gets large.

By the law of total probability, the conditional probability for $\tilde{B} \neq B$ given η_k and ζ_l can now be bounded by

$$\frac{P(\tilde{B}\neq B|\eta)}{P(B|\eta)} \leq \sum_{s=1}^{\infty}\lambda^s n^{-1/2}\bar{q} \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

It follows that

$$\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1 \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

which converges to zero as $n_w \to \infty$. Similarly, exchanging the roles of η and $\tilde{\eta}$, as well as B and $\tilde{B}$, and repeating these steps we can bound

$$\frac{P(B|\eta)}{P(B|\tilde{\eta})} - 1 \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

In order to show that these two inequalities imply the desired bound, we have to distinguish two cases: If $P(B|\tilde{\eta}) \geq P(B|\eta)$, then $\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1 > 0$ so that by first inequality,

$$\left|\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1\right| = \frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1 \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

If on the other hand $P(B|\tilde{\eta}) \leq P(B|\eta)$, then $\frac{P(B|\eta)}{P(B|\tilde{\eta})} - 1 > 0$ so that the second inequality also holds in absolute values. Since in that case we also have

$$\left|\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1\right| \leq \frac{P(B|\eta)}{P(B|\tilde{\eta})}\left|\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1\right| = \left|\frac{P(B|\eta)}{P(B|\tilde{\eta})} - 1\right| \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

Hence the upper bound is the same in both cases, so that

$$\left|\frac{P(B|\tilde{\eta})}{P(B|\eta)} - 1\right| \leq \frac{n^{-1/2}\bar{q}}{1-\lambda}$$

which converges to zero. Combining the two bounds with (B.6) yields the conclusion of part (a).

For parts (b) and (c), note that the argument in part (a) can be extended directly from one to any finite number of individuals. Specifically if we replace η_i and ζ_j with independent copies, we generate four rather than two chains of adjustments, whereas at any iteration, each chain can affect either i's or j's opportunity set. Hence, we can bound the probability of a shift by a multiple of the bound in part (a), $\frac{4n^{-1/2}\bar{q}}{1-\lambda}$, which can in turn be made arbitrarily small by choosing n large enough. Part (c) can be established in a completely analogous fashion. □

In the following, let $I_{wi}^M = I_{wi}[M_i^M]$ and $I_{wi}^W = I_{wi}[M_i^W]$ denote the inclusive values for woman i under the two extremal matchings, so that for any other stable matching, $I_{wi}^M \leq I_{wi}[M_i^*] \leq I_{wi}^W$. Also, let $\Gamma_w^M(x)$ and $\Gamma_w^W(x)$ be the corresponding average inclusive value functions. Similarly, we let $I_{mj}^M = I_{mj}[W_j^M]$ and $I_{mj}^W = I_{mj}[W_j^W]$ the men's inclusive values, and $\Gamma_m^M(z)$ and $\Gamma_m^W(z)$ the corresponding average inclusive value functions.

Lemma B.4. *Suppose Assumptions 2.1, 2.2, and 4.1 hold. Then, (a) for the M-preferred stable matching,*

$$I_{wi}^M \geq \hat{\Gamma}_{wn}^M(x_i) + o_p(1) \text{ and } I_{mj}^M \leq \hat{\Gamma}_{mn}^M(z_j) + o_p(1)$$

for all $i = 1, \dots, n_w$ and $j = 1, \dots, n_m$. Furthermore, (b), if the weight functions $\omega(x,z) \geq 0$ are bounded and form a VC class in (x,z), then

$$\sup_{x\in\mathcal{X}} \frac{1}{n}\sum_{j=1}^{n_m} \omega(x,z_j)(I_{mj}^M - \hat{\Gamma}_m^M(z_j)) \leq o_p(1) \text{ and } \inf_{z\in\mathcal{Z}} \frac{1}{n}\sum_{j=1}^{n_m} \omega(x_i,z)(I_{wi}^M - \hat{\Gamma}_w^M(x_i)) \geq o_p(1).$$

The analogous conclusions hold for the W-preferred stable matching.

Proof: First, note that we can bound conditional choice probabilities given an opportunity set from a pairwise stable matching using the extremal matchings: Specifically, we define

$$\Lambda_w^M(x,z;M^M) := P(U_{ij} \geq U_i^*(M_i^M)|M_i^M = M^M, x_i = x, z_j = z)$$

as the conditional choice probability given the realization of the opportunity set M^M from the male-preferred matching. Also, we let the function $\Lambda_w(x, z, W)$ be the conditional choice probability for an exogenously fixed opportunity set W as defined in (2.4).

By Lemma B.3, the conditional distribution of taste shifters η_i given that W_i^M is supported by a stable matching is approximated by its marginal distribution as n grows large. Hence, combining Lemmas B.1 and B.3,

$$J\Lambda_w^M(x_i, z; M_i^M) \le J\Lambda_w(x_i, z; M_i^M) + o_p(1)$$

Furthermore, the conditional success probabilities $\Lambda_w(\cdot; I_w)$ and $\Lambda_m(\cdot; I_m)$ are of the order $J^{-1} = n^{-1/2}$, whereas the approximation error in Lemma B.3 is multiplicative. Hence, by Lemma B.3 part (b),

$$\mathbb{E}[J(D_{il_1}^M - \Lambda_m(z_{l_1}, x_i; I_{ml_1}))|I_{ml_1}^M, I_{ml_2}^M, x_i, z_j] \to 0,$$

and

$$\mathbb{E}[J^2(D_{il_1}^M - \Lambda_m(z_{l_1}, x_i; I_{ml_1}^M))(D_{il_2}^M - \Lambda_m(z_{l_2}, x_i; I_{ml_2}^M))|I_{ml_1}^M, I_{ml_2}^M, x_i, z_j] \to 0$$

with probability approaching 1, and for all $l_1 = 1, \dots, n_m$ and $l_2 \neq l_1$. Therefore by the law of iterated expectations, and the conditional variance identity we have that for any two men $l_1 \neq l_2$ the unconditional pairwise covariance

$$J^2\mathrm{Cov}((D_{il_1}^M - \Lambda_m(z_{l_1}, x_i; I_{ml_1}^M)), (D_{il_2}^M - \Lambda_m(z_{l_2}, x_i; I_{ml_2}^M))) \to 0$$

with probability approaching 1. Since by Lemma 2.1 $\exp\{U(x, z)\}$ is bounded by a constant, we have that $\mathrm{Var}(I_{wi}^M - \hat{\Gamma}_{wn}^M(x_i)) \to 0$ for each $i = 1, \dots, n_w$, so that part (a) follows from Chebyshev's Inequality.

For part (b), it is sufficient to notice that part (a) and boundedness of $\omega(x, z)$ imply joint convergence in probability for any finite grid of values $x^{(1)}, \dots, x^{(k)} \in \mathcal{X}$, so that uniform convergence follows from the VC condition on $\omega(x, z)$ following standard arguments. □

Next, we establish uniform convergence of the fixed point mapping $\hat{\Psi}$ in equation (4.1). We consider uniformity with respect $\Gamma_w \in \mathcal{T}_w$ and $\Gamma_m \in \mathcal{T}_m$, where $\mathcal{T}_w$ and $\mathcal{T}_m$ denote the space of bounded continuous real-valued functions on $\mathcal{X}$ and $\mathcal{Z}$, respectively, whose values and first p partial derivatives are bounded by constants larger or equal to those from Theorem 3.1.

Recall that $\hat{\Psi}_w[\Gamma_m](x)$ as defined in (4.1) is a sample average

$$\hat{\Psi}_{wn}[\Gamma_m](x) = \frac{1}{n}\sum_{j=1}^{n_m} \psi_w(z_j, x; \Gamma_m)$$

where

$$\psi_w(z_j, x; \Gamma_m) := \frac{\exp\{U(x, z_j; \theta) + V(z_j, x; \theta)\}}{1 + \Gamma_m(z_j)}$$

Similarly, we denote

$$\psi_m(x_i, z; \Gamma_w) := \frac{\exp\{U(x_i, z) + V(z, x_i)\}}{1 + \Gamma_w(x_i)}$$

and define the classes of functions $\mathcal{F}_w : \{\psi_w(\cdot, x; \Gamma_m) : x \in \mathcal{X}, \Gamma_m \in \mathcal{T}_m\}$ and $\mathcal{F}_m : \{\psi_m(\cdot, z; \Gamma_w) : z \in \mathcal{Z}, \Gamma_w \in \mathcal{T}_w\}$.

Lemma B.5. *Suppose Assumption 2.1 holds. Then (i) the classes $\mathcal{F}_w$ and $\mathcal{F}_m$ are Donsker, and (ii) the mapping*

$$\left(\hat{\Psi}_w[\Gamma_m](x), \hat{\Psi}_m[\Gamma_w](x)\right) \stackrel{p}{\to} (\Psi_w[\Gamma_m](x), \Psi_m[\Gamma_w](z))$$

uniformly in $\Gamma_w \in \mathcal{T}_w$ and $\Gamma_m \in \mathcal{T}_m$ and $(x', z')' \in \mathcal{X} \times \mathcal{Z}$ as $n \to \infty$.

Proof: The Donsker property follows from fairly standard arguments: By Assumption 2.1, the function $\exp\{U(x,z;\theta)+V(z,x;\theta)\}$ is Lipschitz continuous in each of its arguments. Following Example 19.7 in van der Vaart (1998), $\mathcal{G} := \{\exp\{U(x,z;\theta)+V(z,x;\theta)\} : x \in \mathcal{X}, z \in \mathcal{Z}, \theta \in \Theta\}$ is a Vapnik-Cervonenkis (VC) class, and therefore also Donsker. Since by definition of $\mathcal{T}_w, \mathcal{T}_m$, $\Gamma_w \in \mathcal{T}_w$ and $\Gamma_m \in \mathcal{T}_m$ have p bounded derivatives, the class $\mathcal{H} = \{\Gamma_w \in \mathcal{T}_w\} \cup \{\Gamma_m \in \mathcal{T}_m\}$ satisfies the conditions for Example 19.9 in van der Vaart (1998), and is also VC. Now note that the transformation $\psi(g,h) := \frac{g}{1+h}$ for $g \in \mathcal{G}$ and $h \in \mathcal{H}$ is continuous and bounded on its domain since g and h are bounded, and $h \geq 0$. It then follows from Example 19.20 in van der Vaart (1998) that the class $\left\{\psi(g,h) := \frac{g}{1+h} | g \in \mathcal{G}, h \in \mathcal{H}\right\}$ is also Donsker.

To establish (ii), note that the Donsker property of $\mathcal{F}_w, \mathcal{F}_m$ implies that the classes are also Glivenko-Cantelli. Hence, $\hat{\Psi}_w$ and $\hat{\Psi}_m$ converge uniformly to their respective population expectations. □

B.3. **Proof of Theorem 4.2:** We now turn to the proof of the main theorem, starting with part (a). Since the theorem assumes the conditions of the Lemmas B.1 and B.3, we have that woman i's conditional probability for man j given the opportunity set M_i converges to $\hat{\Lambda}_w(z_j, x_i; M_i)$. Hence, Lemma B.4 implies convergence of the inclusive value I_{wi} to $\hat{\Gamma}_w(x_i)$,

Fixed-point representation. By Lemma B.4, we have that for the M-preferred matching, $I_{wi}^M \geq \hat{\Gamma}_{wn}^M(x_i) + o_p(1)$ and $I_{mj}^M \leq \hat{\Gamma}_{mn}^M(z_j) + o_p(1)$ for all $i = 1, \dots, n_w$ and $j = 1, \dots, n_m$. Note that by construction $I_{mj}^M \geq 0$ a.s., and $\exp\{U(x,z)+V(z,x)\} \leq \exp\{\bar{U}+\bar{V}\} < \infty$ is bounded by Assumption 2.1, and is a VC class of functions in x, z. Hence we can apply Lemma B.4 part (b) to conclude that

$$\begin{aligned}\hat{\Gamma}_w^M(x) &= \frac{1}{n}\sum_{j=1}^{n_m} \frac{\exp\{U(x,z_j)+V(z_j,x)\}}{1+I_{mj}^M} \\ &\geq \frac{1}{n}\sum_{j=1}^{n_m} \frac{\exp\{U(x,z_j)+V(z_j,x)\}}{1+\hat{\Gamma}_m^M(z_j)} + o_p(1)\end{aligned}$$

where the remainder converges to zero in probability uniformly in x. We obtain similar expressions for $\hat{\Gamma}_m^M(z)$, $\hat{\Gamma}_w^W(x)$, and $\hat{\Gamma}_m^W(z)$. Hence, the inclusive value functions satisfy

$$\begin{aligned}\hat{\Gamma}_w^M &\geq \hat{\Psi}_w^M[\hat{\Gamma}_m^M] + o_p(1) \quad \text{and} \quad \hat{\Gamma}_m^M \leq \hat{\Psi}_m^M[\hat{\Gamma}_w^M] + o_p(1) \\ \hat{\Gamma}_w^W &\leq \hat{\Psi}_w^W[\hat{\Gamma}_m^W] + o_p(1) \quad \text{and} \quad \hat{\Gamma}_m^W \geq \hat{\Psi}_m^W[\hat{\Gamma}_w^W] + o_p(1)\end{aligned}$$

where inequalities are component-wise, i.e. for $\hat{\Gamma}^M(x)$ and $\hat{\Gamma}^W(z)$ evaluated at any value of $x \in \mathcal{X}$ and $z \in \mathcal{Z}$, respectively. Noting that $\hat{\Psi}_w[\Gamma_m]$ and $\hat{\Psi}_m[\Gamma_w]$ are nonincreasing and Lipschitz continuous in Γ_m and Γ_w, respectively, we have

$$\hat{\Gamma}_w^M \geq \hat{\Psi}_w^M[\hat{\Gamma}_m^M] + o_p(1) \geq \hat{\Psi}_w^M[\hat{\Psi}_m^M[\hat{\Gamma}_w^M]] + o_p(1)$$

from the first two inequalities. Hence, for any functions (Γ_w^*, Γ_m^*) solving the fixed-point problem

$$\Gamma_w^* = \hat{\Psi}_w[\Gamma_m^*] + o_p(1) \text{ and } \Gamma_m^* = \hat{\Psi}_m[\Gamma_w^*] + o_p(1)$$

with equality, we have

$$\hat{\Gamma}_w^M \geq \Gamma_w^* + o_p(1) \text{ and } \hat{\Gamma}_m^M \leq \Gamma_m^* + o_p(1)$$

and, from the second set of inequalities,

$$\hat{\Gamma}_w^W \leq \Gamma_w^* + o_p(1) \text{ and } \hat{\Gamma}_m^W \geq \Gamma_m^* + o_p(1)$$

However, since the mapping $\hat{\Psi}$ is a contraction in logs, the fixed point (Γ_w^*, Γ_m^*) is unique up to a term converging to zero in probability. Furthermore, since $M_i^M \subset M_i^W$ and $W_j^W \subset W_j^W$ almost surely, we also have

$$\hat{\Gamma}_w^M \leq \hat{\Gamma}_w^W \text{ and } \hat{\Gamma}_m^M \geq \hat{\Gamma}_m^W$$

It therefore follows that

$$\hat{\Gamma}_w^M = \Gamma_w^* + o_p(1) \text{ and } \hat{\Gamma}_m^M = \Gamma_m^* + o_p(1)$$

and the same condition also holds for the inclusive values from the W-preferred matching. Note that these results need not be uniform with respect to (any random selection from) the full set of stable matchings, but our argument only requires joint convergence for the two extremal matchings.

This establishes the fixed point representation for $\hat{\Gamma}_w^W$ and $\hat{\Gamma}_w^M$ in equations (4.1) and (4.2). Similarly, we can also establish the fixed point characterization for the inclusive value function $\hat{\Gamma}_m^W$ and $\hat{\Gamma}_m^M$ for the male side of the market. Since for any other stable matching, $\hat{\Gamma}_w^M \leq \hat{\Gamma}_w^* \leq \hat{\Gamma}_w^W$ and $\hat{\Gamma}_m^W \leq \hat{\Gamma}_m^* \leq \hat{\Gamma}_m^M$, and furthermore by Theorem 3.1, the solution to the exact fixed-point problem $\Gamma = \hat{\Psi}[\Gamma]$ is unique with probability 1, it follows that (4.2) is also valid for the inclusive value functions under any other stable matching.

In order to prove part (b), we will proceed by the following steps: we first show existence and smoothness of the solutions to the fixed-point problem in the finite economy (4.2), and then show that the solution to the fixed-point problem of the limiting market in (3.5) is well separated, so that uniform convergence of the mapping $\log\hat{\Psi}$ to $\log\Psi$ implies convergence of $\hat{\Gamma}$ to Γ^*.

Existence and smoothness conditions for $\hat{\Gamma}$. First, note that existence and differentiability of $\hat{\Gamma}_w$ and $\hat{\Gamma}_m$ solving the fixed point problem in (4.2) follows from Theorem 3.1: Since the conditions of the theorem do not make any assumptions on the distribution of x_i and z_j, it applies to the case in which $w(x)$ and $m(z)$ are the p.m.f.s corresponding to the empirical distributions of x_i and z_j, respectively. Hence, Assumption 2.1 and Theorem 3.1 imply uniqueness and differentiability to pth order with uniformly bounded partial derivatives conditional on any realization of the empirical distribution of observable characteristics. Since the bounds on the contraction constant λ and on partial derivatives of $\hat{\Gamma}_w, \hat{\Gamma}_m$ do not depend on the marginal distributions, they also hold almost surely with respect to realizations of the empirical distribution.

Local Uniqueness. Next, we verify that for all $\delta > 0$ we can find $\eta > 0$ such that whenever $\tilde{\Gamma}$ such that $\|\log\tilde{\Gamma} - \log\Gamma^*\|_\infty > \delta$ we have $\|(\log\tilde{\Gamma} - \log\Psi[\tilde{\Gamma}]) - (\log\Gamma - \log\Psi[\Gamma])\|_\infty > \eta$: First, note that by Theorem 3.1, the mapping $(\log\Gamma) \mapsto (\log\Psi[\Gamma])$ is a contraction with constant $\lambda := \frac{\exp\{\bar{U}+\bar{V}+\gamma^*\}}{1+\exp\{\bar{U}+\bar{V}+\gamma^*\}} < 1$, where we let $\gamma^* := \max\{\gamma_w, \gamma_m\}$. Then, using the triangle inequality, we can bound

$$\begin{aligned}
\|(\log\tilde{\Gamma} - \log\Psi[\tilde{\Gamma}]) - (\log\Gamma - \log\Psi[\Gamma])\|_\infty &\geq \|\log\tilde{\Gamma} - \log\Gamma^*\|_\infty - \|\log\Psi[\tilde{\Gamma}]) - \log\Psi[\Gamma]\|_\infty \\
&\geq \|\log\tilde{\Gamma} - \log\Gamma^*\|_\infty - \lambda\|\log\tilde{\Gamma} - \log\Gamma^*\|_\infty \\
&> (1-\lambda)\delta > 0
\end{aligned}$$

so that we can choose $\eta = \eta(\delta) := (1-\lambda)\delta$.

Convergence of $\hat{\Gamma} - \Gamma^*$. Finally, Lemma B.5 implies that the fixed point mapping $\hat{\Psi}$ converges to Ψ_0 uniformly in x, z and $\Gamma_w \in \mathcal{T}_w$, and $\Gamma_m \in \mathcal{T}_m$. Since $\hat{\Psi} > 0$ is also bounded away from zero almost surely, it follows that $|\log\hat{\Psi} - \log\Psi_0|$ converges to zero in outer probability and uniformly in in x, z and $\Gamma_w \in \mathcal{T}_w$, and $\Gamma_m \in \mathcal{T}_m$

as well. Hence for any $\varepsilon > 0$ and n large enough, we have

$$P\left(\sup_{\Gamma\in\mathcal{T}} \|\log\hat{\Psi}[\Gamma] - \log\Psi_0[\Gamma]\|_\infty > \frac{\eta}{2}\right) \leq 1-\varepsilon$$

It follows from the choice of η above that

$$P\left(\|\log\hat{\Gamma} - \log\Gamma^*\|_\infty > \delta\right) \leq 1-\varepsilon$$

so that convergence of $\hat{\Gamma}$ to Γ^* in probability under the sup norm follows from the continuous mapping theorem. □

B.4. **Proof of Corollary 4.1.** As shown in section 2, the event that woman i and man j are matched under a stable matching requires that woman i prefers j over any man l in her opportunity set M_i^* given that matching, and that man j prefers i over any woman k in his opportunity set W_j^*. Now, by Lemmata B.1 and B.3 part (a), the conditional probability that i prefers j over any $l \in M_i^*$ given her inclusive value satisfies

$$JP(U_{ij} \geq U_i^*(M_i^*)|I_{wi}, x_i, z_j) = J\Lambda_w(x_i, z_j; I_{wi}) + o(1)$$

with probability approaching 1, where $\Lambda_w(\cdot)$ is as defined in section 2.3. Now, by Theorem 4.2 (b), the inclusive values I_{wi} and I_{mj} converge in probability to $\Gamma_w(x_i)$ and $\Gamma_m(z_j)$, respectively, so that by the continuous mapping theorem,

$$J\Lambda_w(x_i, z_j; I_{wi}) = J\Lambda_w(x_i, z_j; \Gamma_w(x_i)) + o_p(1).$$

Similarly, the conditional probability that man j chooses i over every $k \in W_j^*$ converges according to

$$JP(V_{ji} \geq V_j^*(W_i^*)|I_{mj}, z_j, x_i) = J\Lambda_m(z_j, x_i; \Gamma_m(z_j)) + o_p(1).$$

Finally, by Lemma B.3 part (b) and Assumption 4.1, the joint probability of the two events converges to the product of the marginals,

$$nP(U_{ij} \geq U_i^*(M_i^*), V_{ji} \geq V_j^*(W_i^*)|I_{wi}, I_{mj}, x_i, z_j) = J^2\Lambda_w(x_i, z_j; \Gamma_w(x_i))\Lambda_m(z_j, x_i; \Gamma_m(z_j)),$$

so that the conclusion of this corollary follows from a LLN using B.3 part (c) together with Assumptions 2.1 and 4.1, via an argument analogous to the proof of Lemma B.4. □

References

Agarwal, N. (2012): "An Empirical Model of the Medical Match," working paper, Harvard University.

Azevedo, E., and J. Leshno (2012): "A Supply and Demand Framework for Two-Sided Matching Markets," working paper, Harvard University.

Baccara, M., A. Imrohoroglu, A. Wilson, and L. Yariv (2012): "A Field Study on Matching with Network Externalities," *American Economic Review*, 102(5).

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

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

Christakis, N., J. Fowler, G. Imbens, and K. Kalyanaraman (2010): "An Empirical Model for Strategic Network Formation," working paper, Harvard University.

Cosslett, S. (1988): "Probabilistic Choice over a Continuous Range: An Econometric Model Based on Extreme-Value Stochastic Processes," working paper, Ohio State University.

Dagsvik, J. (1994): "Discrete and Continous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes," *Econometrica*, pp. 1179–1205.

——— (2000): "Aggregation in Matching Markets," *International Economic Review*, 41(1), 27–57.

Decker, C., E. Lieb, R. McCann, and B. Stephens (2013): "Unique Equilibria and Substitution Effects in a Stochastic Model of the Marriage Market," *Journal of Economic Theory*, 148, 778–792.

Dubé, J., J. Fox, and C. Su (2012): "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80(5), 2231–2267.

Dupuy, A., and A. Galichon (2013): "Personality Traits and the Marriage Market," working paper, Sciences-Po.

Echenique, F., S. Lee, and M. Shum (2010): "Aggregate Matchings," working paper, Caltech.

Echenique, F., S. Lee, M. Shum, and B. Yenmez (2012): "The Revealed Preference Theory of Stable and Extremal Stable Matchings," *Econometrica*, forthcoming.

Fox, J. (2010): "Identification in Matching Games," *Quantitative Economics*, 1, 203–254.

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

Galichon, A., and B. Salanié (2010): "Matching with Trade-offs: Revealed Preferences over Competing Characteristics," working paper, Columbia and École Polytechnique.

——— (2012): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," working paper, Columbia and École Polytechnique.

Graham, B. (2013): "Uniqueness, Comparative Static, and Computational Methods for an Empirical One-to-One Transferable Utility Matching Model," working paper, UC Berkeley.

Hall, P. (1979): "On the Rate of Convergence of Normal Extremes," *Journal of Applied Probability*, 16, 433–439.

Hausman, J. (1983): *Specification and Estimation of Simultaneous Equation Models*chap. 7, pp. 391–448. in: Grilliches and Intrilligator (eds.): Handbook of Econometrics, Vol. I.

Hsieh, Y. (2012): "Understanding Mate Preferences from Two-Sided Matching Markets: Identification, Estimation and Policy Analysis," working paper, USC.

Logan, J., P. Hoff, and M. Newton (2008): "Two-Sided Estimation of Mate Preferences for Similarities in Age Education, and Religion," *Journal of the American Statistical Association*, 103(482), 559–569.

McFadden, D. (1974): *Conditional Logit Analysis of Qualitative Choice Behavior*pp. 105–142. Academic Press, New York, in Zarembka (ed.): Frontiers in Econometrics.

Menzel, K. (2012): "Inference for Games with Many Players," working paper, New York University.

Newey, W., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," Handbook of Econometrics, Vol IV Chapter 36.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006): "Moment Inequalities and their Application," working paper, Harvard University.

Resnick, S. (1987): *Extreme Values, Regular Variation, and Point Processes*. Springer.

Resnick, S., and R. Roy (1991): "Random USC Functions, Max-Stable Processes and Continuous Choice," *Journal of Applied Probability*, 1(2), 267–292.

R.Luce (1959): *Individual Choice Behavior*. Wiley, New York.

Roth, A., and M. Sotomayor (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.

Roth, A., and J. Vande Vate (1990): "Random Paths to Stability in Two-Sided Matching," *Econometrica*, 58(6), 1475–1480.

Small, K., and H. Rosen (1981): "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 49(1), 105–130.

Su, C., and K. Judd (2012): "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80(5), 2213–2230.

Uetake, K., and Y. Watanabe (2012): "Entry by Merger: Estimates from a Two-Sided Matching Model with Externality," working paper, Northwestern University.

van der Vaart, A. (1998): *Asymptotic Statistics*. Cambridge University Press, Cambridge.

van der Vaart, A., and J. Wellner (1996): *Weak Convergence and Empirical Processes*. Springer, New York.

Yellott, J. (1977): "The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment and the Double Exponential Distribution," *Journal of Mathematical Psychology*, 15, 109–144.