WORKING PAPERS

N° 1340

June 2022

"Two-Sided Matching Without Transfers: A Unifying Empirical Framework"

Tim Ederer

Toulouse School of Economics

Two-Sided Matching Without Transfers: A Unifying Empirical Framework*

Tim Ederer†

June 15, 2022

Abstract

This paper provides a unifying framework of one-to-one and many-to-one matching without transfers and investigates how data on realized matches can be leveraged to identify preferences of participating agents. I find that, under parsimonious assumptions on preferences, one can only identify the joint surplus function both in the one-to-one and many-to-one case. While this negative identification result was already established for the one-to-one case, I reconcile this finding with the recent literature showing that preferences are separately identified when having data on many-to-one matchings. I find that these positive identification results are mostly driven by restrictions imposed on preferences rather than the additional identification power made available through the many-to-one structure of the data. I then show that by imposing similar restrictions on preferences, one can recover identification of preferences both in the one-to-one and many-to-one case. Finally, I show that the additional data brought by many-to-one matchings can alternatively be used to estimate more precisely the distribution of unobserved preference heterogeneity.

Keywords: two-sided matching, identification, estimation

---

*I sincerely thank Matteo Bobba, Cristina Gualdani and Thierry Magnac for their advice and support. I also thank Nikhil Agarwal and Takuro Yamashita as well as seminar participants in TSE for their useful comments and suggestions.

†Toulouse School of Economics (TSE), Email: tim.ederer@tse-fr.eu

1 Introduction

Two-sided matching models with non-transferable utility are key to understand how centralized clearinghouses allocating jobs, college seats, public housing or deceased-donor kidneys are organized and how one should design them (Roth (2018), Agarwal and Budish (2021)). They are also essential tools to predict the impact of policies aiming at affecting how agents sort in such markets (Agarwal (2017)). However, this requires to know ex-ante the preferences of participating agents which are difficult to infer from observed sorting patterns only. In matching markets, agents' opportunities depend on preferences of agents from the other side. Thus, developing a revealed preference approach based on realized matches is not straightforward in the absence of prior information about preferences of one side of the market.[1]

The goal of this paper is to provide a unifying framework of one-to-one and many-to-one matching without transfers and investigate what can be identified from data on realized matches, when preferences of both sides of the market are unknown. I show that, under parsimonious assumptions on preferences, one can identify the joint surplus function both in the one-to-one and many-to-one case. However, I find that preferences of participating agents cannot be separately identified from the surplus function. Knowing the joint surplus is enough to simulate matching outcomes under various counterfactual scenarios. However, it does not allow us to characterize key objects, such as labor supply elasticities, that solely depend on individual preferences. While this negative identification result was already known for the one-to-one case (Menzel (2015)), it is at odds with the recent literature which highlights that data on many-to-one matching brings additional information that can separately identify preferences (Diamond and Agarwal (2017), He et al. (2021)). This suggests that these positive identification results are mostly driven by other restrictions imposed on preferences, rather than the additional identification power made available through the many-to-one structure of the data. In light of this result, I show that by imposing similar restrictions on preferences one can recover identification of preferences both in the one-to-one and many-to-one case, expanding the scope of what can be learned from data on one-to-one matches. Finally, I show

[1] While such information is sometimes available in college admissions or school choice mechanisms (Agarwal and Somaini (2020)), preferences of both sides of the market are usually unknown.

that the additional data brought by many-to-one matchings can still be useful to estimate more precisely the distribution of unobserved preference heterogeneity.

To perform this analysis, I build on Menzel (2015) to develop a model of two-sided matching where one side is composed of firms and the other side is composed of workers. Each side is characterized by a large set of observed and unobserved attributes. I embed both the one-to-one and many-to-one framework in this model by assuming that each firm has an exogenous finite number of open vacancies, which is larger or equal than one. I impose three assumptions on the payoff functions and the equilibrium: (i) the systematic and unobserved part of the payoff functions are additively separable, (ii) the unobserved taste shocks are *iid* with type-I upper tail and (iii) the observed matching is stable. While (i) and (iii) are commonly used in the literature, (ii) departs from Diamond and Agarwal (2017) and He et al. (2021) by restricting the class of distributions taste shocks can follow for tractability purposes. However, (ii) remains nested in the broader classes they consider implying that the generality of the non identification result derived in this paper is not affected. On the other hand, I do not restrict preferences to be homogeneous and the number of agents on one side of the market to be fixed at the cost of allowing for multiple equilibria.

As in Menzel (2015), I consider that we observe a random sample of realized matches from a single large market where the number of participating firm and workers grows to infinity. Sorting patterns are thus collapsed into the limit of the joint distribution of matched characteristics. Under the assumptions described above, I characterize the mapping between this limit joint distribution function and agents' payoff functions in four steps. First, I show that stability implies that each worker is matched to its preferred firm among the set of firms that would be willing to hire her. Similarly, each firm is matched to its preferred group of workers among the set of workers that would be willing to work there. This implies that we can reinterpret the realized matches as the outcome of two discrete choice models with unobserved and endogenous choice sets, and where firms choose many alternatives. Second, I abstract away from this complexity and derive the limit of workers and firms' conditional choice probabilities under arbitrary exogenous choice sets. Third, I introduce choice sets' endogeneity and show that the information necessary to characterize conditional choice probabilities can be summarized into sufficient statistics called *inclusive values*. Finally, I show that these sufficient statistics converge to the unique solution of a fixed point problem which explicitly links agents' preferences and choice sets. This implies that all stable matches are observationally equivalent and that the limit joint distribution of matched characteristics can be expressed as a function of agents' payoff functions and inclusive values.

By inverting the mapping between the observed sorting and agents' preferences, I find that, without additional data or restrictions on preferences, one can only identify the joint surplus from data on realized matches. This shows that the additional data brought by many-to-one matchings does not help to separately identify agents' preferences. I then show that when the systematic part of the payoff functions is common to all workers/firms (as in Diamond and Agarwal (2017)), one can separately identify preferences from the joint surplus both in the one-to-one case and many-to-one case. Similarly, I find that under appropriate exclusion restrictions (as in He et al. (2021) and Agarwal and Somaini (2022)), one can also recover preferences in the one-to-one case and many-to-one case. I then propose a maximum likelihood estimator that can be tractably used for a parametric version of this framework. Finally, I validate the theoretical limiting results and test the performance of the estimation procedure proposed through Monte Carlo simulations. I find that having data on many-to-one matches allows to estimate more precisely the distribution of random coefficients, mirroring a similar result found for discrete choice models in Berry et al. (2004).

This paper contributes to the literature on empirical models of two-sided matching. One strand of this literature investigates what can be inferred from data on reported preferences within centralized allocation mechanisms (see Agarwal and Somaini (2020) for a review). These methods allowed, for example, to make progress in understanding how school choice mechanisms should be designed (Abdulkadiroğlu et al. (2017), Kapor et al. (2020)). However, in many instances, such data is not available and the econometrician can only rely on realized matches to learn about participating agents' preferences. A large literature examines what can be identified from sorting patterns in models of matching with transferable utility (TU) (Choo and Siow (2006), Fox (2010), Gualdani and Sinha (2019), Galichon and Salanié (Forthcoming)). However, only a handful of papers consider the same problem in the non-transferable utility (NTU) case (see Agarwal and Somaini (Forthcoming) for a review). Menzel (2015) shows that, under parsimonious assumptions on preferences and when matching is one-to-one, only the joint surplus is identified. To circumvent this negative result, Diamond and Agarwal (2017) find that, by restricting preferences to be common to all agents from the same side, one can separately identify preferences with data on many-to-one matches. He et al. (2021) and Agarwal and Somaini (2022) show that, by instead considering a many-to-one matching market where the number of agents on one side is fixed while the other side grows large, exclusion restrictions are sufficient and necessary in order to identify preferences from realized matches. This paper contributes to this literature by providing a unifying empirical framework of one-to-one and many-to-one matching and reconciling the results previously derived in the literature. I find that these recent positive identification results are mostly driven by the extra structure imposed on preferences and not by the inherent additional information brought by having data on many-to-one matches. This means that such methods would also work when having data on one-to-one matches which expands the scope of what can be learned from these models by making them more broadly applicable.[2]

The rest of the paper is organized as follows. Section 2 introduces the preference model along with the equilibrium concept. Section 3 defines the objects that are observed in the data and the sampling process that identifies them. Section 4 establishes the link between the limit joint distribution of matched characteristics and the primitives of the model. Section 5 discusses identification and estimation in the base model, as well as under a various set of additional restrictions on preferences. Section 6 displays results from Monte Carlo simulations.

2 Model

I consider a large two-sided matching market where the number of agents on both sides grows to infinity. I start by introducing the relevant parts of the model in the finite economy before defining the asymptotic sequence that characterizes the limit economy.

Throughout this section, I refer to one side of the market as workers and the other side as firms. Workers are indexed by $i \in \mathcal{I}$ where $\mathcal{I} = \{1, ..., n_w\}$ and firms are indexed by $j \in \mathcal{J}$

[2] In many empirical settings, such as centralized labor clearinghouses or college admissions, firms often open only one vacancy, making Diamond and Agarwal (2017) unapplicable, and the number of agents on both sides of the market is large, making He et al. (2021) and Agarwal and Somaini (2022) unapplicable.

where $\mathcal{J} = \{1, ..., n_m\}$. I nest both the one-to-one and many-to-one matching framework by allowing each firm j to have a finite and exogenous number $q \geq 1$ of open vacancies.[3] I define the matching function μ_w which maps the set of available workers to their matching outcome, which is either their matched employer or the option to remain unmatched. Similarly, μ_m maps the set of available firms to their matching outcome, which is a set of length q including their matched employees as well as the option to leave any open vacancy unfilled.

For instance, consider a given worker i and firm j with $q = 2$. $\mu_w(i) = j$ means that worker i is matched with school j whereas $\mu_m(j) = \{i, l\}$ means that firm j is matched with workers i and l. Similarly, $\mu_w(i) = 0$ means that worker i chooses to stay unmatched, while $\mu_m(j) = \{l, 0\}$ means that firm j is matched with worker l but leaves one of its vacancies unfilled. Note that all elements of the model nest Menzel (2015), which corresponds to the one-to-one case $q = 1$.

2.1 Preferences

Firms and workers are characterized by their observed attributes which collapse into two vectors of random variables $\boldsymbol{x}_i$ and $\boldsymbol{z}_j$. I define their probability distribution functions as $w(\boldsymbol{x})$ and $m(\boldsymbol{z})$ which have support $\mathcal{X}$ and $\mathcal{Z}$, respectively. I specify the utility that worker i gets from being matched with firm j as:

$$U_{ij} = U(\boldsymbol{x}_i, \boldsymbol{z}_j) + \sigma\eta_{ij}$$

whereas the utility that firm j gets from being matched with worker i is defined as:

$$V_{ij} = V(\boldsymbol{x}_i, \boldsymbol{z}_j) + \sigma\epsilon_{ij}$$

ϵ_{ij} and η_{ij} are worker-firm specific unobserved preference shocks and are assumed to be additively separable from the systematic part of the payoffs. I also assume that firms' preferences over groups of workers are responsive (Roth and Sotomayor (1992)). This implies that knowing firms' preferences over individual workers is enough to infer firms' preferences over

[3] Allowing for each firm j to open a different number of vacancies q_j does not affect the main results of the analysis.

groups of workers. Under this assumption, the preferred group of q workers for a given firm is composed of its q individually preferred workers.[4] I impose the following restrictions on the unknown functions U and V and the distribution of unobserved taste shocks.

Assumption 1 *(i). U and V are uniformly bounded in absolute value and $p \geq 1$ times differentiable with uniformly bounded partial derivatives in $\mathcal{X} \times \mathcal{Z}$.*
(ii). ϵ_{ij} and η_{ij} are iid and drawn independently from $\boldsymbol{x}_i$ and $\boldsymbol{z}_j$ from a distribution with absolutely continuous c.d.f. $G(s)$ and density $g(s)$. The upper tail of the distribution $G(s)$ is of type I with auxiliary function $a(s) = \frac{1-G(s)}{g(s)}$.

Assumption 1.(i) is a standard regularity condition which ensures that the functions U and V are well behaved. Assumption 1.(ii) deserves more discussion. It first assumes that observables are independent of unobserved preference shocks. This is usual in discrete choice models but might be particularly strong if we consider a market where prices are set endogenously. However, validity of this assumption can be restored through a control function approach, conditional on having an exogenous price shifter available.[5] Assumption 1.(ii) also imposes restrictions on the upper tail of the distribution of ϵ and η but leaves the lower tail unrestricted. As the number of workers and firms will grow to infinity, the number of independent draws of ϵ and η will also grow. All values of ϵ and η located in the lower tail of their distribution will thus be inconsequential in determining which alternative is the most preferred. As in Menzel (2015), I thus assume that G belongs to a class of distributions which might have different lower tails but for which the upper tail is type I extreme value distributed.[6] Note that this class of functions encompasses most of the parametric distributions traditionally used in discrete choice models. For the Gamma distribution or the Gumbel distribution, this assumption holds for $a(s) = 1$. For the standard normal distribution, this holds for $a(s) = \frac{1}{s}$.

[4] Note that this rules out potential complementarities in preferences over workers. Relaxing this assumption would substantially complicate the analysis given that a stable equilibrium might not even exist in this case.

[5] For example, Agarwal (2015) uses competing hospitals' Medicare reimbursements to instrument for wages in the labor market for medical residents.

[6] This class of distribution is also called the domain of attraction of the Gumbel distribution (Resnick (1987))

2.2 Normalizations

For the limit economy to predict sorting patterns that are consistent with the finite economy, I make several additional assumptions. First, I specify the utility of the outside option as:

$$U_{i0} = \sigma \max_{k=1,\dots,J} \eta_{i0,k}$$

$$V_{0j} = \sigma \max_{k=1,\dots,J} \epsilon_{0j,k}$$

As in Menzel (2015), I then impose the following normalizations on the asymptotic sequence:

Assumption 2 *The asymptotic sequence is controlled by* $n = 1, 2, \dots$ *and we define:*
(i). $n_w = [\exp(\gamma_w)n]$, $n_m = [\exp(\gamma_m)n]$
(ii). $J = [n^{1/2}]$
(iii). $\sigma = \frac{1}{a(b_n)}$ *where* $b_n = G^{-1}(1 - n^{-1/2})$

Assumption 2.(i) allows to control flexibly the relative sizes of each side of the market through the parameters γ_w and γ_m. Even if the size of each side converges to infinity, we can still allow for the mass of agents on one side to be larger than the other side and vice versa. Assumption 2.(ii) makes sure that the probability that workers stay unmatched or that firms keep one vacancy empty does not become degenerate in the limit. If the size of the outside option does not grow with the size of the market, the probability that it becomes dominated by an alternative option will tend to one given that taste shocks have unbounded support. Assumption 2.(iii) controls the scale of the unobserved shocks such that both unobserved and systematic parts of the payoffs jointly determine agents choices in the limit. Given that U and V are bounded and that the support of the taste shocks is unbounded, U and V would become irrelevant in the limit without this restriction. More specifically, if G is Gumbel, then $b_n \asymp \frac{1}{2}\log(n)$ and $\sigma_n = 1$. If taste shocks are standard normal, $b_n \asymp \sqrt{\log n}$ and $\sigma_n \asymp b_n$ and for Gamma distributed taste shocks, $b_n \asymp \log(n)$ and $\sigma_n = 1$.

2.3 Equilibrium

For the remainder of the paper, I refer to a matching as μ which collects μ_m and μ_w and summarizes the matching outcome of each agent. To rationalize the matching we observe and link it to the primitives of our model, I assume that the match is stable.

Definition 1 *For a given $q \geq 1$, a matching μ is stable if and only if for all $i = 1, ..., n_w$ and $j = 1, ..., n_m$:*

(i) Individual rationality: $U_{i\mu_w(i)} \geq U_{i0}$ and $V_{lj} \geq V_{0j}$ for all $l \in \mu_m(j)$.

(ii) No blocking pairs: There exist no pair i, j such that $U_{ij} > U_{i\mu_w(i)}$ and $V_{ij} > \min_{i' \in \mu_m(j)} V_{i'j}$.

A match is stable if agents weakly prefer their match rather than staying unmatched and if there is no worker-firm pair that would prefer be matched together instead of their current match partners. This assumption is typically used in centralized matching markets as it rules out the presence of mismatches due to frictions. Note that, for $q > 1$, this definition is valid only under the assumption that firms' preferences over groups of workers are responsive.[7] Responsiveness also ensures the existence of a stable match and of the worker-optimal/firm-optimal stable matches for $q > 1$ (Roth and Sotomayor (1992)).[8] However, when firms' preferences are heterogeneous, many stable matches can exist and their number grow with the size of the market. I impose no restrictions on which stable outcome is reached in the data. Throughout the rest of the paper, I thus refer to any arbitrary stable match as μ^*. I also define the worker-optimal stable match as μ^W and the firm-optimal stable match as μ^M.

3 Data and Sampling Process

I assume that we observe a sample of realized matches randomly drawn from the limit economy. Observed sorting patterns collapse into the matching frequency distribution function. I define this distribution in the finite economy as the function F_n which gives the expected number of groups of q workers with observable characteristics $(x_1, x_2, ..., x_q)$ matched with firms with observable characteristics z:

$$F_n(x_1, ..., x_q, z; \mu) = \frac{1}{J^{q+1}} \frac{1}{q!} \sum_{i_1=1}^{n_w} ... \sum_{i_q=1}^{n_w} \sum_{j=1}^{n_m} P(x_{i_1} \leq x_1, ..., x_{i_q} \leq x_q, z_j \leq z, \mu_m(j) = \{i_1, ..., i_q\})$$

[7] For $q = 1$, $\mu_m(j)$ is a singleton for all j such that $\min_{i' \in \mu_m(j)} V_{i'j} = V_{\mu_m(j)j}$. We thus recover the same definition as in Menzel (2015).

[8] The worker-optimal stable match is the most preferred stable outcome from the workers' perspective and the least preferred stable outcome from the firms' perspective. On the contrary, the firm-optimal stable match is the most preferred stable outcome from the firms' perspective and the least preferred stable outcome from the workers' perspective.

Normalizing by $q!$ avoids counting the same matched group several times. Alternatively, for firms with observable characteristics z matched to $k < q$ workers with observable characteristics $(x_1, x_2, ..., x_k)$, F_n is defined as:

$$F_n(x_1, ..., x_k, *, z; \mu) =$$
$$\frac{1}{J^{k+1}} \frac{1}{k!} \sum_{i_1=1}^{n_w} ... \sum_{i_k=1}^{n_w} \sum_{j=1}^{n_m} P(x_{i_1} \leq x_1, ..., x_{i_k} \leq x_k, z_j \leq z, \mu_m(j) = \{i_1, ..., i_k\} \cup \{0\}^{q-k})$$

Finally, for firms with observable characteristics z leaving all their vacancies empty and unmatched workers with observable characteristics x, F_n is defined as:

$$F_n(*, z; \mu) = \frac{1}{J^2} \sum_{j=1}^{n_m} P(z_j \leq z, \mu_m(j) = \{0\}^q)$$

$$F_n(x, *; \mu) = \frac{1}{J^2} \sum_{i=1}^{n_w} P(x_i \leq x, \mu_w(i) = 0)$$

I then denote F the limit of the distribution function F_n as the size of the market n grows to infinity. I also define the joint density of matched characteristics f which is the Radon-Nikodym derivative of the limiting measure F.

From there, I link this limiting joint density f to the density of matched characteristics that would arise under various sampling schemes. I assume that the sampling process draws individuals from the population regardless of whether they are firms or workers. One observation is thus composed of this individual alone, if it is unmatched, or along with its matched partners otherwise. Assuming that $q = 1$, the probability that a matched individual is selected by this sampling process is thus twice the probability that an unmatched individual is selected. Indeed, a matched pair could be selected either by drawing the corresponding firm or worker. For any $q \geq 1$, the probability that a matched individual is selected will thus depend on the number of other workers matched to the same firm. Indeed, if a firm is matched with three employees, the probability that any of them is selected is four times the probability that a single agent is selected. I thus define the joint density function arising from this sampling process as:

$$h(x_1, ..., x_q, z) = \frac{(q+1)f(x_1, ..., x_q, z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}}$$

where $h(x_1, ..., x_q, z)$ is the mass of firms with observable z matched with q workers with observed characteristics $(x_1, ..., x_q)$ arising from the sampling scheme defined above and $\exp\{\gamma_w\}+\exp\{\gamma_m\}$ is the total mass of workers and firms available in this economy. Similarly, I define:

$$h(x_1, ..., x_k, *, z) = \frac{(k+1)f(x_1, ..., x_k, *, z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}}$$

$$h(x, *) = \frac{f(x, *)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}}$$

$$h(*, z) = \frac{f(*, z)}{\exp\{\gamma_w\} + \exp\{\gamma_m\}}$$

where $h(x_1, ..., x_k, *, z)$ is the mass of firms with observable z matched with k workers with observed characteristics $(x_1, ..., x_k)$, and $h(x, *)$ with $h(*, z)$ are the mass of unmatched workers and firms. This establishes a direct link between f and h. The next section focuses on linking f with agents' payoff functions.

4 Characterization of the Limit Economy

This section characterizes f, the limiting joint distribution of matched characteristics, as a function of the primitives of the model. The proof follows the same steps as Menzel (2015) and shows how each intermediary result generalizes to $q > 1$. First, I show that stability implies that the realized matches can be interpreted as the outcome of two discrete choice models with endogenous and unobserved choice sets. These choice sets are called opportunity sets and depend on preferences of the other side of the market and which stable match is selected. Second, I consider a simplified economy where opportunity sets would be observed and exogenous and derive the limit of the conditional matching probabilities. Third, I show that, the assumption imposed on the distribution of the tails of the unobserved preference shocks implies that we can use *inclusive values* as sufficient statistics to simplify the problem.

These inclusive values collapse all the information contained in opportunity sets needed to characterize conditional matching probabilities. Finally, I show that these inclusive values can be represented as the approximate solution of a fixed point problem making explicit the relationship between agents' opportunity sets and preferences. This fixed point problem has a unique solution in the limit, which implies that all stable matches are observationally equivalent. I then characterize f as a function of agents' payoff functions and inclusive values.

4.1 Opportunity Sets

Given a match μ^*, I define the opportunity set of a worker as the set of firms that would be willing to hire her instead of one of its current matched employees. Similarly, the opportunity set of a firm is the set of workers that would be willing to quit its current employer to accept a position there. Formally, I define the opportunity set faced by a given worker $i \in \mathcal{I}$ under a match μ^* as:

$$M_i(\mu^*) = \{j \in \mathcal{J} : V_{ij} \geq \min_{i' \in \mu^*_m(j)} V_{i'j}\}$$

Similarly, I define the opportunity set of firm $j \in \mathcal{J}$ as:

$$W_j(\mu^*) = \{i \in \mathcal{I} : U_{ij} \geq U_{i\mu^*_w(i)}\}$$

I then define:

$$U_{i,(k)}(M_i(\mu^*)) = \max\{\min\{U_{ij} : j \in \mathcal{K}\} : \mathcal{K} \subset M_i(\mu^*) \cup \{0\} \text{ and } |\mathcal{K}| = k\}$$

$$V_{j,(k)}(W_j(\mu^*)) = \max\{\min\{V_{ij} : i \in \mathcal{K}\} : \mathcal{K} \subset W_j(\mu^*) \cup \{0\}^k \text{ and } |\mathcal{K}| = k\}$$

where $U_{i,(k)}(M_i(\mu^*))$ denotes the k^{th} highest element of $\{U_{ij'} : j' \in M_i(\mu^*) \cup \{0\}\}$. Note that $U_{i,(1)}(M_i(\mu^*)) = \max_{j' \in M_i(\mu^*) \cup \{0\}} U_{ij'}$. The first important result follows:

Proposition 1 *For any given $q \geq 1$, a match μ^* is stable if and only if for all $i = 1, ..., n_w$*

and $j = 1, ..., n_m$*:*

$$U_{i\mu_w^*(i)} = U_{i,(1)}(M_i(\mu^*)) \quad \textit{and} \quad \forall l \in \mu_m^*(j), V_{lj} \geq V_{j,(q)}(W_j(\mu^*))$$

See Appendix A.1 for a proof of this result. Proposition 1 states that a match μ^* is stable if and only if each worker $i = 1, ..., n_w$ is matched to her preferred alternative among her opportunity set and each firm $j = 1, ..., n_m$ is matched to the q^{th} highest ranked alternatives among its opportunity set. This implies the following corollary:

Corollary 1 *For a given stable match* μ^* *and any worker* i *and firm* j*:*
(i). $j = \mu_w^*(i) \iff i \in \mu_m^*(j) \iff U_{ij} \geq U_{i,(1)}(M_i(\mu^*)) \quad \textit{and} \quad V_{ij} \geq V_{j,(q)}(W_j(\mu^*))$
(ii). $0 \in \mu_m^*(j) \iff V_{0j} \geq V_{j,(q)}(W_j(\mu^*))$
(iii). $\mu_w^*(i) = 0 \iff U_{i0} \geq U_{j,(1)}(M_i(\mu^*))$

This corollary states that a stable match μ^* can be rewritten as the outcome of two discrete choice models where each agent's choice set is its opportunity set. This equivalence establishes a link between the observed matching and the primitives of the model. However, opportunity sets are unobserved and endogenous objects as they depend on μ^* and on the preferences of agents from the other side of the market. Additionally, characterizing the probability of being among a given firm's q^{th} most preferred workers is not standard when $q > 1$. Deriving the limit of conditional matching probabilities is thus not straightforward.

4.2 Limit of Conditional Choice Probabilities

To simplify the analysis, I consider here arbitrary exogenous opportunity sets $M_i = \{1, ..., J\}$ and $W_j = \{1, ..., J\}$. From Corollary 1, we know that conditional matching probabilities can be characterized as two-sided conditional choice probabilities:

$$\begin{aligned}\mathbb{P}(j = \mu_w(i)|x_i, z_j) &= \mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i) \quad \text{and} \quad V_{ij} \geq V_{j,(q)}(W_j)|x_i, z_j) \\ &= \mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i)|x_i, z_j) \times \mathbb{P}(V_{ij} \geq V_{j,(q)}(W_j)|x_i, z_j)\end{aligned}$$

The limit of these conditional choice probabilities have the following expression:

Proposition 2 *Under Assumption 1 and 2, as $J \to \infty$ for a given finite $q \geq 1$ and for all i and j:*

$$J\mathbb{P}(U_{ij} \geq U_{i,(q)}(M_i)|x_i, z_j) \longrightarrow \exp(U(x_i, z_j)) \times \left[1 - \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds}\right)^q\right]$$

$$\mathbb{P}(U_{i0} \geq U_{i,(q)}(M_i)|x_i) \longrightarrow \left[1 - \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds}\right)^q\right]$$

See Appendix A.2 for a proof of this result. Proposition 2 generalizes the Logit formula to cases where agents make many unranked choices among an infinite number of alternatives. For $q = 1$, we recover the usual Logit formula derived in Menzel (2015). Note that the Independence of Irrelevant Alternatives (IIA) property still holds for $q > 1$. This has several implications regarding how one can allow for realistic substitution patterns. Imposing the distribution of ϵ and η to be normal would not change this result as the normal distribution has a type-I upper tail. To avoid this problem, one can alternatively introduce unobserved preference heterogeneity for observed characteristics through the use of random coefficients. Note that the CCP of choosing a particular alternative j would converge to zero if we do not weight it by J, the rate at which the total number of alternatives increases. Lemma 1 in Appendix A.3 establishes that the size of opportunity sets increases at a rate $\sqrt{n}$ which justifies Assumption 2.(ii).

4.3 Inclusive Values

I now introduce that opportunity sets are actually endogenous and unobserved. Endogeneity arises as shifting worker i's taste shocks could make her prefer another feasible firm to its current match. This could then trigger a chain of rematches that could potentially affect her own opportunity set. This problem is even more salient in the context of many-to-one matching as changing firm j's taste shocks could trigger at most q chains of rematches, which increases the probability that this ends up changing firm j's opportunity set. However, as in Menzel (2015), I find that, as the size of the market increases, the probability for such an event to occur vanishes to zero. This result stems mostly from two implications of Proposition 2: (i) the probability that firm j rematches with a specific worker i vanishes to zero as the size of opportunity sets increase to infinity and (ii) the probability of choosing the outside option instead, which would terminate such a chain of rematches, is non degenerate in the limit. This result is formalized in Lemma 2 in Appendix A.3 where a more detailed discussion and proof can be found.

From this, I then show that the dependence between taste shocks and opportunity sets vanishes in the limit. This means that the distribution of taste shocks conditional on opportunity sets converges to their marginal distribution g. However, this claim can only be proven for the opportunity sets derived from the extremal matchings. The distribution of taste shocks conditional on opportunity sets is only well defined for the extremal matchings, given that they are the only stable matchings that always exist irrespective of the size of the market. Again, this result is formalized in Lemma 3 in Appendix A.3. This means that we can use this result along with Proposition 2 to bound[9] the CCPs conditional on the opportunity sets that would arise under the firm-optimal stable match μ^M as follows:

$$\begin{aligned} & n^{1/2}\mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^M))|x_i, z_j, (z_k)_{k\in M_i(\mu^M)}, M_i(\mu^M)) \\ & \leq \frac{\exp\{U(x_i, z_j)\}}{1 + n^{-1/2}\sum_{k\in M_i(\mu^M)}\exp\{U(x_i, z_k)\}} + o(1) \end{aligned} \tag{1}$$

$$\begin{aligned} & n^{1/2}\mathbb{P}(V_{ij} \geq V_{j,(q)}(W_j(\mu^M))|x_i, z_j, (x_l)_{l\in W_j(\mu^M)}, W_j(\mu^M)) \\ & \geq \exp(V(x_i, z_j)) \times \left[1 - \left(\frac{n^{-1/2}\sum_{l\in W_j(\mu^M)}\exp\{V(x_l, z_j)\}}{1 + n^{-1/2}\sum_{l\in W_j(\mu^M)}\exp\{V(x_l, z_j)\}}\right)^q\right] + o(1) \end{aligned} \tag{2}$$

Similar bounds can be computed for the worker-optimal stable match μ^W where the direction of the inequalities is reversed. In Equation 1 and 2, $n^{-1/2}\sum_{k\in M_i(\mu^W)}\exp\{U(x_i, z_k)\}$ and $n^{-1/2}\sum_{l\in W_j(\mu^M)}\exp\{V(x_l, z_j)\}$ serve as sufficient statistics that collapse all the information contained in opportunity sets which is needed to approximate CCPs. These objects are called *inclusive values*.

[9] Note that we only provide bounds given that there are several potential stable matches μ^* such that $M_i(\mu^*) = M_i(\mu^M)$ and $W_j(\mu^*) = W_j(\mu^M)$.

More generally, I define worker i's inclusive value given a realized stable match μ^* as:

$$I^*_{wi} = n^{-1/2} \sum_{j \in M_i(\mu^*)} \exp(U(x_i, z_j))$$

Similarly, I define firm j's inclusive value given μ^* as:

$$I^*_{mj} = n^{-1/2} \sum_{i \in W_j(\mu^*)} \exp(V(x_i, z_j))$$

I also define I^M_{wi} and I^M_{mj} as the inclusive values that would arise under the firm-optimal stable match and I^W_{wi} and I^W_{mj} as the inclusive values that would arise under the worker-optimal stable match.

Of course, in practice, inclusive values are unobserved and we do not know which stable match is selected. The rest of this section shows that the inclusive values arising from any stable match μ^* can be approximated by the solution of a fixed point problem which has a unique solution in the limit.

4.4 Fixed Point Characterization for Inclusive Values

I first show that, for any $q \geq 1$, inclusive values arising from the firm-optimal and worker-optimal stable match can be approximated by *expected inclusive value functions* (Menzel (2015)). I first rewrite I^M_{wi} as:

$$\begin{aligned} I^M_{wi} &= \frac{1}{n} \sum_{k=1}^{n_m} \exp\{U(x_i, z_k)\} \times \sqrt{n} \mathbb{1}\{k \in M_i(\mu^M)\} \\ &= \frac{1}{n} \sum_{k=1}^{n_m} \exp\{U(x_i, z_k)\} \times \sqrt{n} \mathbb{1}\{V_{ik} \geq V_{k,(q)}(W_k(\mu^M))\} \end{aligned}$$

The inclusive value of a given worker is determined by the set of firms that would accept her, which in turn depends on the preferences of all firms as well as their opportunity sets. Using Equation 2, I then show that:

$$I^M_{wi} \geq \hat{\Gamma}^M_w(x_i) + o_p(1)$$

where $\hat{\Gamma}_w^M$ is the firm-optimal *expected inclusive value function* of workers which is defined as:

$$\hat{\Gamma}_w^M(x_i) = \frac{1}{n}\sum_{k=1}^{n_m} \exp\{U(x_i, z_k) + V(x_i, z_k)\} \times \left[1 - \left(\frac{I_{mk}^M}{1 + I_{mk}^M}\right)^q\right]$$

Similarly, using Equation 1, I show that we can approximate I_{mj}^M as follows:

$$I_{mj}^M \leq \hat{\Gamma}_m^M(z_j) + o_p(1)$$

where $\hat{\Gamma}_m^M$ is the firm-optimal *expected inclusive value function* of firms which is defined as:

$$\hat{\Gamma}_m^M(z_j) = \frac{1}{n}\sum_{l=1}^{n_w} \frac{\exp\{U(x_l, z_j) + V(x_l, z_j)\}}{1 + I_{wl}^M}$$

Note that similar bounds can be established for the inclusive values that would arise under the worker-optimal stable match:

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

A formal exposition and proof of this result can be found in Lemma 4 in Appendix A.3. The inclusive value of a given worker can be approximated by a function of firms' preferences and inclusive values. Similarly, the inclusive value of a given firm can be approximated by a function of workers' preferences and inclusive values. Hence, the two-sided nature of the problem gives rise naturally to a fixed point problem characterizing these inclusive values. I define the fixed point mappings as follows:

$$\hat{\Psi}_w[\Gamma_m](x) = \frac{1}{n}\sum_{k=1}^{n_m} \exp\{U(x, z_k) + V(x, z_k)\} \times \left[1 - \left(\frac{\Gamma_m(z_k)}{1 + \Gamma_m(z_k)}\right)^q\right]$$

$$\hat{\Psi}_m[\Gamma_w](z) = \frac{1}{n}\sum_{l=1}^{n_w} \frac{\exp\{U(x_l, z) + V(x_l, z)\}}{1 + \Gamma_w(x_l)}$$

From there, I show, using Lemma 4, that for any $x \in \mathcal{X}$ and $z \in \mathcal{Z}$:

$$\hat{\Gamma}_w^M(x) \geq \hat{\Psi}_w[\hat{\Gamma}_m^M](x) + o_p(1) \quad \text{and} \quad \hat{\Gamma}_m^M(z) \leq \hat{\Psi}_m[\hat{\Gamma}_w^M](z) + o_p(1) \tag{3}$$

$$\hat{\Gamma}_w^W(x) \leq \hat{\Psi}_w[\hat{\Gamma}_m^W](x) + o_p(1) \quad \text{and} \quad \hat{\Gamma}_m^W(z) \geq \hat{\Psi}_m[\hat{\Gamma}_w^W](z) + o_p(1) \tag{4}$$

In addition, the firm-optimal stable match is unanimously preferred by firms while the worker-optimal stable match is unanimously preferred by workers (Roth and Sotomayor (1992)). This implies that $M_i(\mu^M) \subset M_i(\mu^*) \subset M_i(\mu^W)$ and $W_i(\mu^W) \subset W_i(\mu^*) \subset W_i(\mu^M)$ which means that for all i and j:

$$I_{wi}^M \leq I_{wi}^* \leq I_{wi}^W \quad \text{and} \quad I_{mj}^W \leq I_{mj}^* \leq I_{mj}^M$$

This in turn implies that for all (x, z):

$$\hat{\Gamma}_w^M(x) \leq \hat{\Gamma}_w^*(x) \leq \hat{\Gamma}_w^W(x) \quad \text{and} \quad \hat{\Gamma}_m^W(z) \leq \hat{\Gamma}_m^*(z) \leq \hat{\Gamma}_m^M(z)$$

Using Equation 3 and 4, we can thus show that, for any stable matching μ^*:

$$\hat{\Gamma}_w^*(x) = \hat{\Psi}_w[\hat{\Gamma}_m^*](x) + o_p(1) \quad \text{and} \quad \hat{\Gamma}_m^*(z) = \hat{\Psi}_m[\hat{\Gamma}_w^*](z) + o_p(1) \tag{5}$$

which concludes the proof that inclusive values arising from any given stable match μ^* can be approximated by the solution of a fixed point problem.

I now introduce the population equivalent of the fixed point problem described in Equation 5:

$$\Gamma_w^* = \Psi_w[\Gamma_m^*] \quad \text{and} \quad \Gamma_m^* = \Psi_m[\Gamma_w^*] \tag{6}$$

where

$$\Psi_w[\Gamma_m](x) = \int \exp(U(x,s) + V(x,s) + \gamma_m) \times \left[1 - \left(\frac{\Gamma_m(s)}{1+\Gamma_m(s)}\right)^q\right] m(s)ds$$

$$\Psi_m[\Gamma_w](z) = \int \frac{\exp(U(s,z) + V(s,z) + \gamma_w)}{1+\Gamma_w(s)} w(s)ds$$

This population fixed point problem has a unique solution and the approximate solution of the finite sample fixed point problem converges to it. This is stated in the following result:

Theorem 1 *Under Assumption 1 and 2 and for any* $q \geq 1$*:*

(i). The mapping $(\log \Gamma_w, \log \Gamma_w) \mapsto (\log \Psi_m[\Gamma_w], \log \Psi_w[\Gamma_m])$ *is a contraction.*

(ii). The fixed point problem described in Equation 6 always has a unique solution Γ_m^*, Γ_w^**.*

(iii). For any given μ^*, $I^*_{wi} \longrightarrow \Gamma^*_w(x_i)$ *and* $I^*_{mj} \longrightarrow \Gamma^*_m(z_j)$ *for all* i *and* j.

A proof of this result can be found in Appendix A.3. Theorem 1 has several implications. First, it implies that for any $q \geq 1$ and for any arbitrary stable match μ^*, inclusive values converge to the same limit. This means that all stable matches are observationally equivalent in the limit both in the one-to-one and many-to-one case. This implies that we do not need to have any information about the equilibrium selection mechanism nor do we need to impose restrictions on preferences to ensure that there is a unique stable match[10] to infer preferences from observed sorting. Second, it implies that, for any $q \geq 1$, we can use inclusive value functions as sufficient statistics to characterize CCPs as functions which only depend on agents' observable characteristics. Additionally, we know that the fixed point mappings are contractions which means that solving for inclusive value functions is computationally feasible.

4.5 Limit of Distribution of Matched Characteristics

Finally, using Theorem 1, I characterize the limit of the conditional matching probabilities as follows.

Proposition 3 *(i) For any firm* j *with* $q \geq 1$ *vacancies and any group of workers* $i = 1, ..., k$ *where* $k < q$*:*

$$J^{k+1}\mathbb{P}(\mu_m(j) = \{1, ..., k\} \cup \{0\}^{q-k} | (x_i)_{i=1}^k, z_j) \longrightarrow \frac{k! \exp\left\{\sum_{i=1}^k U(x_i, z_j) + V(x_i, z_j)\right\}}{\prod_{i=1}^k (1 + \Gamma^*_w(x_i)) (1 + \Gamma^*_m(z_j))^{k+1}}$$

(ii) For any firm j *with* $q \geq 1$ *vacancies:*

$$J^{q+1}\mathbb{P}(\mu_m(j) = \{1, ..., q\} | (x_i)_{i=1}^q, z_j) \longrightarrow \frac{q! \exp\left\{\sum_{i=1}^q U(x_i, z_j) + V(x_i, z_j)\right\}}{\prod_{i=1}^q (1 + \Gamma^*_w(x_i))(1 + \Gamma^*_m(z_j))^q}$$

$$\mathbb{P}(\mu_m(j) = \{0\}^q | z_j) \longrightarrow \frac{1}{1 + \Gamma^*_m(z_j)}$$

[10] Assuming that firms' preferences are homogenous, as in Diamond and Agarwal (2017), makes the stable match unique (Roth and Sotomayor (1992)). Similarly, when we assume that there is a continuum of students matching with a fixed number of colleges, as in He et al. (2021) and Agarwal and Somaini (2022), there exists a unique stable match (Azevedo and Leshno (2016)).

A proof of this result can be found in Appendix A.4. The probability that a given match is formed is thus positively correlated with the total match surplus $\sum_{i=1}^{q} U(x_i, z_j) + V(x_i, z_j)$. However, it is negatively correlated with inclusive values as they grow with the size of the set of other potential matching opportunities. It can also be noted that the rate at which these quantities converge to their limits depend on q. The larger is q the slower convergence is. This introduces a trade-off as increasing q might bring additional identification power at the cost of introducing bias due to approximation errors. From this, I characterize the limit joint distribution of matched characteristics:

$$f(x_1, ..., x_k, *, z) = \frac{\exp\left\{ \sum_{l=1}^{k} U(x_l, z) + V(x_l, z) + k\gamma_w + \gamma_m \right\}}{\Pi_{l=1}^{k}(1 + \Gamma_w^*(x_l))(1 + \Gamma_m^*(z))^{k+1}} m(z)\Pi_{l=1}^{k} w(x_l)$$

$$f(x_1, ..., x_q, z) = \frac{\exp\left\{ \sum_{l=1}^{q} U(x_l, z) + V(x_l, z) + q\gamma_w + \gamma_m \right\}}{\Pi_{l=1}^{q}(1 + \Gamma_w^*(x_l))(1 + \Gamma_m^*(z))^{q}} m(z)\Pi_{l=1}^{q} w(x_l)$$

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

$$f(*, z) = \frac{\exp(\gamma_m) m(z)}{1 + \Gamma_m^*(z)}$$

Where $f(x_1, ..., x_k, *, z)$ is the mass of firms with observable z matched with k workers with characteristics $(x_1, ..., x_k)$, $f(x_1, ..., x_q, z)$ is the mass of firms with observable z matched with q workers with characteristics $(x_1, ..., x_q)$, $f(x, *)$ is the mass of unmatched workers with characteristic x and $f(*, z)$ is the mass of unmatched firms with observable z.

5 Identification and Estimation

5.1 Identification Joint Surplus

From the expression of f derived in the previous section, we can show that:

$$\frac{f(x_1, ..., x_k, *, z)}{f(x_1, ..., x_{k-1}, *, z)} = \frac{\exp\left\{U(x_k, z) + V(x_k, z) + \gamma_w\right\}}{(1 + \Gamma_w^*(x_k))(1 + \Gamma_m^*(z))} w(x_k)$$

Inverting this mapping finally gives us:

$$U(x_k, z)+V(x_k, z) = \log f(x_1, ..., x_k, *, z)-\log f(x_1, ..., x_{k-1}, *, z)-\log f(x_k, *)-\log \frac{f(*, z)}{\exp(\gamma_m)m(z)}$$

Given that we can identify f directly from the data, as was discussed in Section 3, this implies that we can identify the surplus function $U + V$. Similarly, we can express inclusive values as functions of the distribution of the characteristics of unmatched individuals:

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

$$\Gamma^*_m(z) = \frac{\exp(\gamma_m)m(z)}{f(*, z)} - 1$$

However, we cannot express U as a function of f separately from V and vice versa. This result is formalized in the following proposition.

Proposition 4 *Under Assumption 1 and 2 and for any $q \geq 1$:*
(i) The joint surplus function $U+V$ and the inclusive value functions Γ^*_w and Γ^*_m are identified from the limiting joint distribution of matched characteristics f.
(ii) Without further restrictions, we cannot separately identify U and V.

This means that the additional data available when $q > 1$ does not bring any additional information which would be useful to separately identify individual preferences from the joint surplus. This is in sharp contrast with Diamond and Agarwal (2017) and He et al. (2021) which find that preferences can be separately identified with data on many-to-one matching. This suggests that these positive identification results mostly rely on the extra assumptions they impose on preferences rather than the additional information made available by the many-to-one structure of the data. This would mean that, by using similar restrictions, we could thus achieve similar positive identification results even for $q = 1$. The goal of the remainder of this section is to verify this claim.

5.2 Homogeneous preferences

I mimic the framework developed in Diamond and Agarwal (2017) by assuming that the systematic part of the payoff functions is homogeneous across individuals. I thus define the utility that worker i gets from being matched with school j as:

$$u_{ij} = U(\boldsymbol{z}_j) + \sigma\eta_{ij}$$

whereas the utility that firm j gets from being matched with worker i is defined as:

$$v_{ij} = V(\boldsymbol{x}_i) + \sigma\epsilon_{ij}$$

Additionally, I assume that there exists $\bar{\boldsymbol{x}}$ such that $V(\bar{\boldsymbol{x}}) = 0$. Note that this framework differs from Diamond and Agarwal (2017) on two dimensions. Taste shocks are heterogeneous and *iid* over i, j and the class of distribution to which they belong is more restrictive. Under these assumptions, it is immediate to see that we can recover U and V from the joint surplus as $U(\boldsymbol{z}) + V(\bar{\boldsymbol{x}}) = U(\boldsymbol{z})$. I state the following result:

Proposition 5 *Under Assumptions 1 and 2 and for any $q \geq 1$, the payoff functions U and V are identified from the limiting joint distribution of matched characteristics f.*

This shows that a similar positive identification result as the one derived in Diamond and Agarwal (2017) can actually be achieved for both $q > 1$ and $q = 1$ by using similar restrictions on preferences. In fact, this suggests that their non identification result for $q = 1$ is mostly driven by the assumption they impose on the correlation structure of the unobserved taste shocks. As is pointed out by the authors, assuming that taste shocks are common to all agents from the same side makes the unique stable match perfectly assortative along these unobserved tastes. This creates an endogeneity problem. It thus becomes necessary to have data on at least two-to-one matching in order to have an additional measurement of these sorting patterns that would allow to disentangle the effect of observed and unobserved preferences. In the framework developed in this paper, this problem does not exist given that taste shocks are *iid* across individuals. In the limit, conditional matching probabilities are uniquely determined by observable characteristics even when $q = 1$.

5.3 Exclusion restrictions

As in He et al. (2021) and Agarwal and Somaini (2022), I assume that a set of variables affecting the utility of one side can be excluded from the utility of the other side. I define the utility that worker i gets from being matched with firm j as:

$$U_{ij} = U(\boldsymbol{x}_i, \boldsymbol{z}_j) + \sigma \eta_{ij}$$

whereas the utility that firm j gets from being matched with worker i is defined as:

$$V_{ij} = V(\boldsymbol{x}_i, \boldsymbol{z}_j) + g(w_i) + \sigma \epsilon_{ij}$$

I additionally assume that g is increasing in w and that $\lim_{w \to \infty} g(w) = \infty$.[11] I also assume that there exists $\bar{w}$ such that $g(\bar{w}) = 0$. Under these assumptions, we can state the following result:

Proposition 6 *Under Assumptions 1 and 2 and for any $q \geq 1$, the payoff functions U, V and g are identified from the limiting joint distribution of matched characteristics f.*

A proof of this result can be found in Appendix A.5. Similarly to the argument used in He et al. (2021) and Agarwal and Somaini (2022), increasing w shifts the probability that a given firm becomes available which allows us to disentangle the role of firms' and workers' preferences in determining the sorting patterns we observe.[12] This argument also holds for $q = 1$ and the many-to-one structure of the data does not help in making this additional source of identification more salient. Note that we do not need here to have preference shifters for both sides of the market as in Agarwal and Somaini (2022) and He et al. (2021). As the joint surplus is already identified in the absence of exclusion restrictions, we only need to identify preferences of workers to recover preferences of firms from the surplus.[13]

[11]This is similar to Assumption 2 in Agarwal and Somaini (2022).

[12]The identification argument only works at infinity as the match has not a fixed cutoff structure as in He et al. (2021) and Agarwal and Somaini (2022). As both the number of firms and workers grow to infinity in our case, the cutoffs grow to infinity as the size of the market grows.

[13]In Agarwal and Somaini (2022) and He et al. (2021) it is not clear whether the joint surplus is identified in the absence of exclusion restrictions. Further work could determine whether it is only the case when taste shocks have type-I upper tails.

5.4 Unobserved Preference Heterogeneity

In light of the previous results, one can wonder how we could use the additional information made available by data on many-to-one matching, if not for disentangling preferences from the joint surplus. Using a similar argument as what is used in the discrete choice literature, I claim that having data on several decisions made by the same firm allows to know more about the unobserved "type" this firm belongs to. More specifically, if we were to assume that firms have an unobserved individual and heterogeneous taste for a given worker characteristic x_i, having more than one measurement of given firm j's choice would be useful to pin it down. This is analogous to what is argued by Berry et al. (2004) who show that estimating random coefficients from a cross section of observed choices often fails when having only the first ranked choice of each consumer. Having at least the second choice of each consumer allows to disentangle what drives observed choices between random coefficients and unobserved taste shocks. A similar argument could apply with data on many-to-one matching given that we observe several workers matched to the same firm. I investigate in Section 6, whether such gains could also be achieved in a matching market setting thanks to the many-to-one structure of the data.

5.5 Estimation

Given that the identification proof is constructive, one could construct naturally a non-parametric estimator for the joint surplus function $U + V$ and the inclusive value functions Γ_w^* and Γ_m^*. However, this would quickly become intractable as the dimensionality of x and z increases.

I instead consider a parametric version of this framework where I define the payoff functions as $U(x, z; \boldsymbol{\theta})$ and $V(x, z; \boldsymbol{\theta})$. I assume that U and V are known for all (x, z) up to a vector of unknown parameters $\boldsymbol{\theta}$. Assume that we observe a random sample of K individuals, drawn from the sampling scheme described in Section 3, along with their respective matches. For a given observation k, we observe a vector $(x_1(k), ..., x_q(k), z(k))$ which has a different structure depending on the type of match we observe. For an unmatched worker, which is indexed by $w(k) = 0$, I record its characteristics in $x_1(k)$ and encode the other variables as missing. For an unmatched firm, which is indexed by $m(k) = 0$, I record its characteristics in $z(k)$. And for a firm matched with a group of workers of size n, indexed by $m(k) = n$, I record the characteristics of all matched workers along with the characteristic of the matched firm in $(x_1(k), ..., x_n(k), z(k))$ and encode the rest as missing. We can then construct the following sample average log-likelihood:

$$
\begin{aligned}
L(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\theta}) = \frac{1}{K} \sum_{k=1}^{K} & \mathbb{1}\{w(k) = 0\} h(x(k), *, \boldsymbol{\theta}) + \mathbb{1}\{m(k) = 0\} h(*, z(k), \boldsymbol{\theta}) \\
& + \mathbb{1}\{m(k) = 1\} h(x_1(k), *, z(k), \boldsymbol{\theta}) \\
& + \mathbb{1}\{m(k) = 2\} h(x_1(k), x_2(k), *, z(k), \boldsymbol{\theta}) \\
& + ... \\
& + \mathbb{1}\{m(k) = q\} h(x_1(k), ..., x_q(k), z(k), \boldsymbol{\theta})
\end{aligned}
$$

where h is the joint density of matched characteristics under the sampling scheme described in Section 3. Of course, calculating the likelihood function for a given parameter $\boldsymbol{\theta}$ first involves solving for the fixed point problem described in Equation 6 to derive the inclusive value functions. This can be achieved by setting up an inner loop that will apply the contraction mapping until convergence. The estimator proposed is then defined as:

$$
\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta}{\arg\max}\, L(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{\theta})
$$

Asymptotic inference for $\hat{\boldsymbol{\theta}}$ is then standard as long as the size of the sample is not too large relative to the size of the overall economy. As is pointed out in Menzel (2015) and Diamond and Agarwal (2017), the inherent structure of matching markets could introduce dependence between observations. A bootstrap procedure could then be used for inference (Diamond and Agarwal (2017), Menzel (2021)).

6 Monte Carlo Simulations

In this section, I perform several Monte Carlo simulations in order to assess: (i) the validity of the convergence results derived in Section 4 and (ii) the validity of the estimation strategy

Table 1: Monte Carlo: Convergence of Matching Frequencies and Inclusive Values

n	Unmatched Workers	Firms with One Unfilled Vacancy	Firms with Two Unfilled Vacancies	I_w	I_m
20	0.5894	0.6778	0.2339	0.7617	0.5343
50	0.5714	0.6712	0.2290	0.8007	0.5365
100	0.5608	0.6661	0.2285	0.8245	0.5350
200	0.5521	0.6621	0.2279	0.8406	0.5344
500	0.5449	0.6588	0.2273	0.8546	0.5313
1000	0.5418	0.6577	0.2263	0.8611	0.5320
2000	0.5389	0.6561	0.2268	0.8662	0.5335
Model	0.5321	0.6527	0.2267	0.8794	0.5321

NOTES. This table reports the average share of unmatched firms and workers in each period taken over 200 sample draws for different sample sizes n.

proposed in Section 5.

6.1 Convergence of Conditional Match Probabilities

Consider a simple model where $q = 2$ and $U(\boldsymbol{x}, \boldsymbol{z}) = V(\boldsymbol{x}, \boldsymbol{z}) = 0$ for all $(x, z) \in \mathcal{X} \times \mathcal{Z}$. In this example, we can easily solve for the fixed point problem described in Equation 6 given that inclusive value functions collapse to a fixed number which does not vary with $(\boldsymbol{x}, \boldsymbol{z})$. This results in $\Gamma_w^* = 0.8794$ and $\Gamma_m^* = 0.5321$. I also compute the limit matching frequencies:

$$\mathbb{P}\left(U_{i0} \geq U_{i,(1)}^*\right) \longrightarrow 0.5321$$

$$\mathbb{P}\left(V_{0j} \geq V_{i,(1)}^*\right) \longrightarrow 0.2267$$

$$\mathbb{P}\left(V_{i,(1)}^* > V_{0j} \geq V_{i,(2)}^*\right) \longrightarrow 0.6527$$

To verify the validity of the large market approximation, I first simulate n individuals along with their taste shocks over the individuals from the other side of the market ϵ_{ij} and η_{ij} for all (i, j). I then use the worker-proposing Deferred Acceptance algorithm to get the worker-optimal stable match. Finally, I compute the empirical matching frequencies and the inclusive values under this stable match and check whether they converge to their theoretical limits as n grows large. Table 1 displays the result of this exercise. Both the inclusive values and the matching frequencies converge to their theoretical limits. This table also shows that

Table 2: Monte Carlo: Estimation without Random Coefficient

	$q=1$			$q=2$			$q=3$		
n	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$
100	0.989	1.048	0.511	0.942	1.053	0.494	0.944	1.051	0.457
	(0.309)	(0.374)	(0.231)	(0.202)	(0.334)	(0.173)	(0.159)	(0.372)	(0.152)
200	0.999	0.962	0.487	0.957	1.047	0.486	0.951	1.045	0.468
	(0.223)	(0.236)	(0.150)	(0.155)	(0.229)	(0.121)	(0.125)	(0.236)	(0.113)
500	0.991	0.997	0.497	0.976	1.018	0.479	0.966	1.013	0.476
	(0.153)	(0.155)	(0.101)	(0.101)	(0.153)	(0.082)	(0.082)	(0.143)	(0.071)
1000	0.989	0.992	0.500	0.987	1.009	0.486	0.974	1.005	0.485
	(0.108)	(0.111)	(0.069)	(0.066)	(0.102)	(0.055)	(0.060)	(0.097)	(0.047)
Model	1	1	0.5	1	1	0.5	1	1	0.5

NOTES. This table reports the average and standard deviation of the ML estimator of $\boldsymbol{\theta}$ for different values of q taken over 200 sample draws for different sample sizes n.

the limit economy is a relatively good approximation even when the size of the market is moderately large.

6.2 Convergence of ML Estimator

I now evaluate the performance of the estimator proposed in Section 5 through two Monte Carlo exercises. In the first exercise, I consider the following simple parametric framework:

$$U_{ij} = \theta_1 z_j + \eta_{ij} \quad \text{and} \quad V_{ij} = \theta_2 x_i + \theta_3 x_i z_j + \epsilon_{ij}$$

and estimate $\boldsymbol{\theta}$ on simulated data. To do so, I draw n individuals along with their observed characteristics x_i and z_j drawn from a standard normal distribution. I draw the taste shocks ϵ_{ij} and η_{ij} from the Gumbel distribution and set $\boldsymbol{\theta} = (1, 1, 0.5)$ to compute U_{ij} and V_{ij} for all (i, j). I then derive the worker-optimal stable match using the Deferred Acceptance algorithm. Finally, I estimate $\boldsymbol{\theta}$ and repeat this process 200 times to report the mean and standard deviation of $\hat{\boldsymbol{\theta}}$ over the sample draws. Table 2 shows that the estimator seems to converge to its true value as the size of the market increases given that the mean converges to the true value while the standard deviation vanishes. We can also see that, while increasing q lowers the variance of the estimator, it also seems to introduce bias. This is consistent with Proposition 3, given that the joint conditional matching probabilities converge to their theoretical limits at a slower rate when q increases. There is thus a trade off involved as

Table 3: Monte Carlo: Estimation with Random Coefficient

	$q=1$				$q=2$				$q=3$			
n	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\mathbb{P}(\hat{\theta}_3=0)$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\mathbb{P}(\hat{\theta}_3=0)$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\mathbb{P}(\hat{\theta}_3=0)$
100	1.006	0.997	0.300	0.485	0.995	0.992	0.311	0.280	0.976	0.990	0.267	0.315
	(0.334)	(0.352)	(0.363)	-	(0.217)	(0.307)	(0.260)	-	(0.168)	(0.265)	(0.234)	-
200	1.037	0.969	0.286	0.425	0.993	0.965	0.317	0.210	0.964	0.960	0.327	0.110
	(0.191)	(0.206)	(0.301)	-	(0.152)	(0.208)	(0.224)	-	(0.113)	(0.195)	(0.169)	-
500	0.995	0.974	0.294	0.310	0.992	0.970	0.350	0.065	0.991	0.977	0.389	0.025
	(0.125)	(0.130)	(0.256)	-	(0.086)	(0.122)	(0.146)	-	(0.068)	(0.128)	(0.116)	-
1000	1.007	0.963	0.348	0.115	0.995	0.981	0.405	0.005	0.995	0.981	0.414	0
	(0.100)	(0.101)	(0.190)	-	(0.060)	(0.086)	(0.105)	-	(0.056)	(0.077)	(0.073)	-
Model	1	1	0.5	-	1	1	0.5	-	1	1	0.5	-

NOTES. This table reports the average and standard deviation of the ML estimator of $\boldsymbol{\theta}$ for different values of q taken over 200 sample draws for different sample sizes n.

increasing q allows to have a more precise estimator, given that we are using more information, but might introduce distortions as conditional matching frequencies converge to their limit at a slower rate.

In a second exercise, I now consider the following parametric framework:

$$U_{ij} = \theta_1 z_j + \eta_{ij} \quad \text{and} \quad V_{ij} = \theta_2 x_i + \theta_3 x_i \nu_j + \epsilon_{ij}$$

where ν_j is unobserved and follows a $\mathcal{N}(0,1)$. In this example, I assume that there is unobserved heterogeneity in schools' tastes over x_i which is parametrized through a normal distributed random coefficient with mean θ_2 and standard deviation θ_3. I then follow similar steps as for the first exercise to get the mean and the standard deviation of the ML estimator of $\boldsymbol{\theta}$ for different values of q and different sizes of the economy n. Note that to approximate the integral over ν_i to compute the conditional matching probabilities, I use a Gaussian-Hermite quadrature (Judd (1998)). Table 3 shows that $\hat{\boldsymbol{\theta}}$ converges to its true value as n increases. However, the standard deviation of the random coefficient θ_3 is poorly estimated when $q=1$. Even with large n, $\hat{\theta}_3$ is equal to 0 in 11% of the cases. Given that the log-likelihood function is symmetric around $\theta_3 = 0$, this indicates that it is maximized while being not differentiable at this point. In this case, traditional inference breaks down and this estimator is not informative. Although this issue also arises for $q=2$ and $q=3$ when n is small, we can see that as q increases this is less likely to happen. In fact, increasing

q from 1 to 2 is already enough to drastically reduce the probability of estimating θ_3 to 0. This indicates that having data on two-to-one matching is already enough to bring additional identification power necessary to pin down the distribution of random coefficients. This mirrors the result found in Berry et al. (2004) which shows that having data on consumers' second choices allows to estimate random coefficients more easily.

7 Conclusion

This paper develops a unifying empirical framework of one-to-one and many-to-one matching without transfers to understand what can be inferred on agents' preferences from observed sorting in such markets. I impose few restrictions on preferences and assume that the observed matching is stable. Stability allows me to rewrite the model as a two-sided discrete choice model with endogenous and unobserved choice sets. I use a sufficient statistics approach to take into account choice sets' endogeneity and characterize agents' conditional choice probabilities. This allows me to form a clear mapping between the joint distribution of matched characteristics and agents' payoff functions.

I then show that we can identify the joint surplus from both one-to-one and many-to-one matching data. However, without further restrictions, individual preferences are not identified. While this negative identification result was already established in the one-to-one case, the literature has argued that many-to-one matchings can bring additional information which would allow to separately identify preferences from the joint surplus. I find that these positive identification results are ultimately not driven by the availability of such additional information but mostly by the extra assumptions imposed on preferences.

I then argue that, by imposing similar restrictions on preferences, one can extend these positive identification result to the one-to-one matching case. More specifically, by either assuming that the systematic parts of the payoffs is homogenous across individuals (as in Diamond and Agarwal (2017)) or under appropriate exclusion restrictions (as in He et al. (2021) and Agarwal and Somaini (2022)), one can separately identify preferences from the joint surplus both in the one-to-one and many-to-one case. Finally, I show that the additional information brought by the many-to-one structure of the data can instead be used to estimate more precisely the distribution of random coefficients in a parametric framework.

A Proofs

A.1 Proof of Proposition 1

Suppose first that μ is not stable. This could imply first, by definition of stability, that there exists a pair (i, j) such that $U_{ij} > U_{i\mu_w(i)}$ and $V_{ij} > \min_{i' \in \mu_m(j)} V_{i'j}$. This would mean that there exists a pair (i, j) such that $j \in M_i(\mu)$ and $U_{ij} > U_{i\mu_w(i)}$ which contradicts that $U_{i\mu_w(i)} = \max_{j' \in M_i(\mu) \cup \{0\}} U_{ij'}$. This could also imply that $U_{i0} > U_{i\mu_w(i)}$ or $V_{0j} > \min_{i' \in \mu_m(j)} V_{i'j}$. In the first case, this would contradict that $U_{i\mu_w(i)} = \max_{j' \in M_i(\mu) \cup \{0\}} U_{ij'}$. In the second case, this would mean that there exist a $l \in \mu_m(j)$ such that $V_{0j} > V_{lj}$ which contradicts that $V_{lj} \geq V_{j,(q)}(W_j(\mu))$.

Now, suppose that for a given i, $U_{i\mu_w(i)} < \max_{j' \in M_i(\mu) \cup \{0\}} U_{ij'}$. This means that there exist a firm $k \in M_i(\mu) \cup \{0\}$ such that $U_{ik} > U_{i\mu_w(i)}$. If $k = 0$, this immediately contradicts stability. If $k \in M_i(\mu)$, this implies that there exist a firm k such that $V_{ik} \geq \min_{i' \in \mu_m(k)} V_{i'k}$ and $U_{ik} > U_{i\mu_w(i)}$. If $V_{ik} = \min_{i' \in \mu_m(k)} V_{i'k}$, this implies that $k = \mu_w(i)$ and we reach a contradiction. Otherwise, we have that $U_{ik} > U_{i\mu_w(i)}$ and $V_{ik} > \min_{i' \in \mu_m(k)} V_{i'k}$ which contradicts stability.

Finally, suppose that for a given j and for a given $l \in \mu_m(j)$, $V_{lj} < V_{j,(q)}(W_j(\mu))$. This means that there exist a worker s such that $s \in W_j(\mu) \cup \{0\}$ and $V_{sj} > V_{lj}$. If $s = 0$, this contradicts stability. If $s \in W_j(\mu)$, this implies that $U_{sj} \geq U_{s\mu_w(s)}$ and $V_{sj} > V_{lj}$. Again, we restrict ourselves to the case where $j \neq \mu_w(s)$ which implies that $U_{sj} > U_{s\mu_w(s)}$ and $V_{sj} > V_{lj}$ which contradicts stability. This concludes the proof.

A.2 Proof of Proposition 2

I first consider the case $q = 2$. The proof for $q = 1$ can be found in Menzel (2015). I start by decomposing in two terms the conditional probability that U_{ij} is above or equal $U_{i,(2)}(M_i)$ where $M_i = \{0, ..., J\}$. I remove the dependence on M_i for simplicity such that $U_{i,(q)}(M_i) = U_{i,(q)}$ for all q. I also rewrite $U_{ij} = u_{ij} + \sigma\eta_{ij}$ for simplicity.

$$\mathbb{P}(U_{ij} \geq U_{i,(2)} | (u_{ik})_{k=1}^J) = \mathbb{P}(U_{ij} \geq U_{i,(1)} | (u_{ik})_{k=1}^J) + \mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)} | (u_{ik})_{k=1}^J)$$

The first term is known already and is the conditional choice probability for $q = 1$. The second term can be expressed as the probability that there exists one alternative preferred to j but that j is preferred to the rest:

$$\begin{aligned}
\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)} | (u_{ik})_{k=1}^J) &= \int \sum_{k=1}^J \mathbb{P}(U_{ik} > U_{ij}, U_{ij} \geq U_{il}, l \in \mathcal{I} - \{k, j\} | (u_{ik})_{k=1}^J, \eta_{ij} = s) g(s) ds \\
&= \int \sum_{k=1}^J (1 - G(\sigma^{-1}(u_{ij} - u_{ik}) + s)) \prod_{l \in \mathcal{I} - \{k,j\}} G(\sigma^{-1}(u_{ij} - u_{il}) + s) g(s) ds \\
&= \int \sum_{k=1}^J \frac{1 - G(\sigma^{-1}(u_{ij} - u_{ik}) + s)}{G(\sigma^{-1}(u_{ij} - u_{ik}) + s)} \prod_{l=1}^{2J} G(\sigma^{-1}(u_{ij} - u_{il}) + s) \frac{g(s)}{G(s)} ds
\end{aligned}$$

As in Menzel (2015), I then do the change of variables $s = a_J t + b_J$ where $a_J = a(b_J)$ and $b_J = G^{-1}(1 - J^{-1/2})$ and multiply by J on both sides:

$$\begin{aligned}
J\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)} | (u_{ik})_{k=1}^J) = &\int \frac{1}{J} \sum_{k=1}^J \frac{J(1 - G(a_J(u_{ij} - u_{ik} + t) + b_J))}{G(a_J(u_{ij} - u_{ik} + t) + b_J)} \\
&\times \exp\left(\frac{1}{J} \sum_{l=1}^{2J} J \log G(a_J(u_{ij} - u_{il} + t) + b_J) \right) \frac{J a_J g(a_J t + b_J)}{G(a_J t + b_J)} dt
\end{aligned}$$

Following Resnick (1987) and under Assumption 1 we can show that:

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

$$G(a_J(u_{ij} - u_{ik} + t) + b_J) \to 1$$

$$J \log G(a_J(u_{ij} - u_{il} + t) + b_J) \to -e^{-(u_{ij} - u_{ik} + t)}$$

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

We thus have under Assumption 1:

$$J\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)}|(u_{ik})_{k=1}^J) = \int \frac{1}{J}\sum_{k=1}^J e^{-(u_{ij}-u_{ik}+t)} \exp\left(-\frac{1}{J}\sum_{l=1}^{2J} e^{-(u_{ij}-u_{ik}+t)}\right) e^{-t}dt + o(1)$$
$$= \int \frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})} \exp\left(-\frac{1}{J}\sum_{l=1}^{2J} e^{-t}e^{(u_{ik}-u_{ij})}\right) e^{-t}e^{-t}dt + o(1)$$

I then do a final change of variable $s = e^{-t}$ such that we get:

$$J\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)}|(u_{ik})_{k=1}^J) = \int_0^{+\infty} \frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})} \exp\left(-\frac{1}{J}\sum_{l=1}^{2J} se^{(u_{ik}-u_{ij})}\right) sds + o(1)$$
$$= \frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})} \left(\frac{1}{J}\sum_{k=1}^{2J} e^{(u_{ik}-u_{ij})}\right)^{-2} + o(1)$$
$$= \frac{1}{J}\sum_{k=1}^J e^{u_{ik}} \frac{\exp(u_{ij})}{\left(\frac{1}{J}\sum_{k=1}^{2J}\exp(u_{ik})\right)^2} + o(1)$$
$$= \frac{\exp(u_{ij})}{1+\frac{1}{J}\sum_{k=1}^J \exp(u_{ik})} \times \frac{\frac{1}{J}\sum_{k=1}^J \exp(u_{ik})}{1+\frac{1}{J}\sum_{k=1}^J \exp(u_{ik})} + o(1)$$

From this we can finally show that:

$$J\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)}|x_i,(z_k)_{k=1}^J) = \frac{\exp(U(x_i,z_j))}{1+\frac{1}{J}\sum_{k=1}^J \exp(U(x_i,z_k))} \times \frac{\frac{1}{J}\sum_{k=1}^J \exp(U(x_i,z_k))}{1+\frac{1}{J}\sum_{k=1}^J \exp(U(x_i,z_k))} + o(1)$$

which implies that:

$$J\mathbb{P}(U_{i,(1)} > U_{ij} \geq U_{i,(2)}|x_i,z_j) \longrightarrow \frac{\exp(U(x_i,z_j))}{1+\int \exp(U(x_i,s))m(s)ds} \times \frac{\int \exp(U(x_i,s))m(s)ds}{1+\int \exp(U(x_i,s))m(s)ds}$$

We know from Menzel (2015) that:

$$J\mathbb{P}(U_{ij} \geq U_{i,(1)}|x_i,z_j) \longrightarrow \frac{\exp(U(x_i,z_j))}{1+\int \exp(U(x_i,s))m(s)ds}$$

So we can conlude that:

$$J\mathbb{P}(U_{ij} \geq U_{i,(2)}|x_i,z_j) \longrightarrow \frac{\exp(U(x_i,z_j))}{1+\int \exp(U(x_i,s))m(s)ds} \times \left(1+\frac{\int \exp(U(x_i,s))m(s)ds}{1+\int \exp(U(x_i,s))m(s)ds}\right)$$

Following similar steps, we can prove that:

$$\mathbb{P}(U_{i0} \geq U_{i,(2)}|x_i) \longrightarrow \frac{1}{1+\int \exp(U(x_i,s))m(s)ds} \times \left(1 + \frac{\int \exp(U(x_i,s))m(s)ds}{1+\int \exp(U(x_i,s))m(s)ds}\right)$$

To illustrate how this iterates to any q, I write here the proof for $q = 3$. Similarly, we want to characterize the probability that there exists two alternatives preferred to j but that j is preferred to the rest:

$$\mathbb{P}(U_{i,(2)} > U_{ij} \geq U_{i,(3)}|(u_{ik})_{k=1}^J)$$

$$= \int \frac{1}{2}\sum_{k=1}^J \sum_{\substack{m=1\\m\neq k}}^J \mathbb{P}(U_{ik} > U_{ij}, U_{im} > U_{ij}, U_{ij} \geq U_{il}, l \in \mathcal{I} \setminus \{k,m,j\}|(u_{ik})_{k=1}^J, \eta_{ij} = s)f(s)ds$$

$$= \int \frac{1}{2}\sum_{k=1}^J \frac{1-G(\sigma^{-1}(u_{ij}-u_{ik})+s)}{G(\sigma^{-1}(u_{ij}-u_{ik})+s)} \sum_{\substack{m=1\\m\neq k}}^J \frac{1-G(\sigma^{-1}(u_{ij}-u_{im})+s)}{G(\sigma^{-1}(u_{ij}-u_{im})+s)} \prod_{l=1}^{2J} G(\sigma^{-1}(u_{ij}-u_{im})+s)\frac{g(s)}{G(s)}ds$$

$$= \int \frac{1}{2}\frac{1}{J}\sum_{k=1}^J \frac{J(1-G(\sigma^{-1}(u_{ij}-u_{ik})+s))}{G(\sigma^{-1}(u_{ij}-u_{ik})+s)} \frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^J \frac{(J-1)(1-G(\sigma^{-1}(u_{ij}-u_{im})+s))}{G(\sigma^{-1}(u_{ij}-u_{im})+s)}$$

$$\times \exp\left(\frac{1}{J}\sum_{l=1}^J J\log G(\sigma^{-1}(u_{ij}-u_{il})+s)\right)\frac{Jg(s)}{G(s)}ds$$

$$= \int \frac{1}{2}\frac{1}{J}\sum_{k=1}^J e^{-(u_{ij}+t-u_{ik})}\frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^J e^{-(u_{ij}+t-u_{im})} \exp\left(-\frac{1}{J}\sum_{l=1}^{2J}\exp -(u_{ij}+s-u_{il})\right)e^{-t}dt + o(1)$$

$$= \int \frac{1}{2}\frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})}\frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^J e^{(u_{im}-u_{ij})} \exp\left(-e^{-t}\frac{1}{J}\sum_{l=1}^{2J}\exp(u_{il}-u_{ij})\right)e^{-t}e^{-t}e^{-t}dt + o(1)$$

$$= \int_0^{+\infty} \frac{1}{2}\frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})}\frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^J e^{(u_{im}-u_{ij})} \exp\left(-s\frac{1}{J}\sum_{l=1}^{2J}\exp(u_{il}-u_{ij})\right)s^2ds + o(1)$$

$$= \frac{1}{J}\sum_{k=1}^J e^{(u_{ik}-u_{ij})}\frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^J e^{(u_{im}-u_{ij})} \times \left(\frac{1}{J}\sum_{l=1}^{2J}\exp(u_{il}-u_{ij})\right)^{-3} + o(1)$$

$$= \frac{\exp(u_{ij})}{\frac{1}{J}\sum_{k=1}^{2J}\exp(u_{ik})} \times \frac{\frac{1}{J}\sum_{k=1}^{J}\exp(u_{ik})\frac{1}{J-1}\sum_{\substack{m=1\\m\neq k}}^{J}\exp(u_{im})}{\left(\frac{1}{J}\sum_{l=1}^{2J}\exp(u_{il})\right)^2} + o(1)$$

$$= \frac{\exp(u_{ij})}{1+\frac{1}{J}\sum_{k=1}^{J}\exp(u_{ik})} \times \left(\frac{\frac{1}{J}\sum_{k=1}^{J}\exp(u_{ik})}{1+\frac{1}{J}\sum_{k=1}^{J}\exp(u_{ik})}\right)^2 + o(1)$$

From this we then have that:

$$J\mathbb{P}(U_{i,(2)} > U_{ij} \geq U_{i,(3)}|x_i, z_j) \longrightarrow \frac{\exp(U(x_i, z_j))}{1+\int \exp(U(x_i, s))m(s)ds} \times \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1+\int \exp(U(x_i, s))m(s)ds}\right)^2$$

We can thus conclude that:

$$J\mathbb{P}(U_{ij} \geq U_{i,(3)}|x_i, z_j) \longrightarrow \frac{\exp(U(x_i, z_j))}{1+\int \exp(U(x_i, s))m(s)ds} \times \sum_{k=1}^{3}\left(\frac{\int \exp(U(x_i, s))m(s)ds}{1+\int \exp(U(x_i, s))m(s)ds}\right)^{k-1}$$

$$\mathbb{P}(U_{i0} \geq U_{i,(3)}|x_i, z_j) \longrightarrow \frac{1}{1+\int \exp(U(x_i, s))m(s)ds} \times \sum_{k=1}^{3}\left(\frac{\int \exp(U(x_i, s))m(s)ds}{1+\int \exp(U(x_i, s))m(s)ds}\right)^{k-1}$$

To prove this result for any q, I derive the limit of the following conditional probability:

$$\mathbb{P}(U_{i,(q-1)} > U_{ij} \geq U_{i,(q)}|(u_{ik})_{k=1}^{J})$$

Following the same steps, the probability that there exists $q-1$ alternatives preferred to j but that j is preferred to the rest can be expressed as:

$$\mathbb{P}(U_{i,(q-1)} > U_{ij} \geq U_{i,(q)}|(u_{ik})_{k=1}^{J})$$
$$= \int \frac{1}{(q-1)!}\sum_{j_1=1}^{J} \dots \sum_{\substack{j_{q-1}=1\\ j_{q-1}\notin\{j_1,\dots,j_{q-2}\}}}^{J} \mathbb{P}(U_{ij_1} > U_{ij}, \dots, U_{ij_{q-1}} > U_{ij}, U_{ij} \geq U_{il}, l \in \mathcal{I}\setminus\{j_1, \dots, j_{q-1}, j\}|(u_{ik})_{k=1}^{J}, \eta_{ij} = s)f(s)ds$$

which results in:

$$J\mathbb{P}(U_{i,(q-1)} > U_{ij} \geq U_{i,(q)}|x_i, z_j) \longrightarrow \frac{\exp(U(x_i, z_j))}{1+\int \exp(U(x_i, s))m(s)ds} \times \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1+\int \exp(U(x_i, s))m(s)ds}\right)^{q-1}$$

We can thus derive the following result:

$$J\mathbb{P}(U_{ij} \geq U_{i,(q)}|x_i, z_j) \longrightarrow \frac{\exp(U(x_i, z_j))}{1 + \int \exp(U(x_i, s))m(s)ds} \times \sum_{k=1}^{q} \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds} \right)^{k-1}$$
$$= \exp(U(x_i, z_j)) \times \left[1 - \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds} \right)^q \right]$$

$$\mathbb{P}(U_{i0} \geq U_{i,(q)}|x_i, z_j) \longrightarrow \frac{1}{1 + \int \exp(U(x_i, s))m(s)ds} \times \sum_{k=1}^{q} \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds} \right)^{k-1}$$
$$= \left[1 - \left(\frac{\int \exp(U(x_i, s))m(s)ds}{1 + \int \exp(U(x_i, s))m(s)ds} \right)^q \right]$$

This concludes the proof of Proposition 2.

A.3 Proof of Theorem 1

I start by proving part (i) and (ii) of Theorem 1. As in Menzel (2015), I first restrict the space of functions to which the solutions to the fixed point problem described in Equation 6 can belong. Namely, I show that we can restrict ourselves to a Banach space of continuous functions.

Assume that there exists a pair of functions Γ_w^* and Γ_m^* that solve the fixed point problem. By definition of Ψ_w and using that $\Gamma_m^* \geq 0$, we have for any $q \geq 1$:

$$\Gamma_w^*(x) = \Psi[\Gamma_m^*](x) = \int \exp\{U(x, s) + V(x, s)\} \left[1 - \left(\frac{\Gamma_m^*(s)}{1 + \Gamma_m^*(s)} \right)^q \right] m(s)ds$$
$$\leq \int \exp\{U(x, s) + V(x, s)\} m(s)ds$$
$$\leq \exp\{\bar{U} + \bar{V}\}$$

where $\bar{U}$ and $\bar{V}$ are the upper bounds of the functions U and V, respectively. Given that this bound holds also for $q = 1$, this implies that we can bound similarly $\Gamma_m^*(x) \leq \exp\{\bar{U} + \bar{V}\}$.

I now establish continuity of the solutions Γ_w^* and Γ_m^*. By definition, I rewrite:

$$\begin{aligned}\Gamma_w^*(x) &= \Psi_w[\Psi_m[\Gamma_w^*]](x)\\ &= \int \exp\{U(x,s)+V(x,s)\}\left[1-\left(\frac{\int \frac{\exp\{U(t,s)+V(t,s)\}}{1+\Gamma_w^*(t)}w(t)dt}{1+\int \frac{\exp\{U(t,s)+V(t,s)\}}{1+\Gamma_w^*(t)}w(t)dt}\right)^q\right]m(s)ds\end{aligned}$$

Similarly, I write:

$$\begin{aligned}\Gamma_m^*(z) &= \Psi_m[\Psi_w[\Gamma_m^*]](z)\\ &= \int \frac{\exp\{U(t,z)+V(t,z)\}}{1+\int \exp\{U(t,s)+V(t,s)\}\left[1-\left(\frac{\Gamma_m^*(s)}{1+\Gamma_m^*(s)}\right)^q\right]m(s)ds}w(t)dt\end{aligned}$$

Since U and V are continuous and all the integrals are nonnegative, $\Psi_w[\Psi_m[\Gamma_w^*]]$ and $\Psi_m[\Psi_w[\Gamma_m^*]]$ are also continuous which establishes continuity of the solutions Γ_w^* and Γ_m^*. Differentiability of Γ_w^* and Γ_m^* also follows from differentiability of U and V which is stated in Assumption 1. We can thus restrict the spaces in which Γ_w^* and Γ_m^* belong to a Banach space of nonnegative bounded continuous functions that I call $\mathcal{C}^*$.

Consider now two pairs of functions (Γ_w,Γ_m) and $(\tilde{\Gamma}_w,\tilde{\Gamma}_m)$ belonging to $\mathcal{C}^*\times\mathcal{C}^*$. I first rewrite:

$$\log\Psi_w[\log\Gamma_m](x) = \int \exp\{U(x,s)+V(x,s)\}\left[1-\left(\frac{\exp\{\log\Gamma_m^*(s)\}}{1+\exp\{\log\Gamma_m^*(s)\}}\right)^q\right]m(s)ds$$

Given that Ψ_w and Ψ_m are Gâteaux differentiable, I use the mean value inequality to establish that:

$$\begin{aligned}&\Big|\Big|\log\Psi_w[\Gamma_m](x)-\log\Psi_w[\tilde{\Gamma}_m](x)\Big|\Big|_\infty\\ &\qquad\leq \sup_{a\in[0,1]}\Big|\Big|d\log\Psi_w[a\log\Gamma_m+(1-a)\log\tilde{\Gamma}_m](x)\Big|\Big|_\infty\Big|\Big|\log\Gamma_m(x)-\log\tilde{\Gamma}_m(x)\Big|\Big|_\infty\end{aligned}$$

where we can write:

$$d\log\Psi_w[\log\Gamma_m](x) = -\frac{1}{\Psi_w[\log\Gamma_m](x)}\int \frac{\exp\{U(x,s)+V(x,s)\}}{1+\Gamma_m(s)}\left(\frac{\Gamma_m(s)}{1+\Gamma_m(s)}\right)^q m(s)ds$$

Rearranging this expression gives the following:

$$d \log \Psi_w[\log \Gamma_m](x) = -\frac{1}{\Psi_w[\log \Gamma_m](x)} \int \exp\{U(x,s) + V(x,s)\} \left[1 - \left(\frac{\Gamma_m(s)}{1+\Gamma_m(s)}\right)^q\right] \times \frac{q(\Gamma_m(s))^q}{(1+\Gamma_m(s))^{q+1} - (\Gamma_m(s))^q - (\Gamma_m^*(s))^{q+1}} m(s)ds$$

Since Γ_m^* has to be positive, we can show that:

$$\begin{aligned}\frac{q(\Gamma_m(s))^q}{(1+\Gamma_m(s))^{q+1} - (\Gamma_m(s))^q - (\Gamma_m(s))^{q+1}} &= \frac{q(\Gamma_m(s))^q}{\sum_{k=0}^{q+1} \frac{(q+1)!}{k!(q+1-k)!}(\Gamma_m(s))^k - (\Gamma_m(s))^q - (\Gamma_m(s))^{q+1}} \\ &= \frac{q(\Gamma_m(s))^q}{\sum_{k=0}^{q-1} \frac{(q+1)!}{k!(q+1-k)!}(\Gamma_m(s))^k + q(\Gamma_m(s))^q} \\ &\leq \frac{q(\exp\{\bar{U}+\bar{V}\})^q}{\sum_{k=0}^{q-1} \frac{(q+1)!}{k!(q+1-k)!}(\exp\{\bar{U}+\bar{V}\})^k + q(\exp\{\bar{U}+\bar{V}\})^q} := \lambda_q\end{aligned}$$

This implies that:

$$\begin{aligned}|d \log \Psi_w[\log \Gamma_m](x)| &\leq \frac{\lambda_q}{\Psi_w[\log \Gamma_m](x)} \int \exp\{U(x,s) + V(x,s)\} \left[1 - \left(\frac{\Gamma_m^*(s)}{1+\Gamma_m^*(s)}\right)^q\right] m(s)ds \\ &= \lambda_q \leq 1\end{aligned}$$

From this, I conclude that:

$$\sup_{a \in [0,1]} \left|\left| d \log \Psi_w[a \log \Gamma_m + (1-a) \log \tilde{\Gamma}_m](x) \right|\right|_\infty \leq \lambda_q$$

which implies that:

$$\left|\left| \log \Psi_w[\Gamma_m](x) - \log \Psi_w[\tilde{\Gamma}_m](x) \right|\right|_\infty \leq \lambda_q \left|\left| \log \Gamma_m(x) - \log \tilde{\Gamma}_m(x) \right|\right|_\infty$$

Given that this holds for any $q \geq 1$, I conclude that the mapping $(\log \Gamma_w, \log \Gamma_m) \mapsto (\log \Psi_w[\Gamma_m], \log \Psi_m[\Gamma_w])$ is a contraction which proves claim (i) of Theorem 1. The proof of part (ii) is a direct implication of the Banach fixed point theorem.

Before proving part (iii) of Theorem 1, intermediary steps are needed. In what follows, I first prove that the size of opportunity sets grow at a rate $\sqrt{n}$ for any $q \geq 1$. From this, I then show that the dependence between opportunity sets and taste shocks under the extremal matchings vanishes as n grows to infinity. I then use this result to show that we can approximate inclusive values arising from any stable match by inclusive value functions which have an approximate fixed point representation. I then finally prove that the solution to the finite sample fixed point problem converges to the unique solution of the population fixed point problem which concludes the proof of Theorem 1.(iii).

A.3.1 Rate of Size of Feasible Choice Sets

Define, for a given stable matching μ^*, the number of firms feasible to worker i and the number of workers feasible to firm j as:

$$J^*_{wi} = \sum_{j=1}^{n_m} \mathbb{1}\{V_{ji} \geq V_{j,(q)}(W_j(\mu^*))\} \quad \text{and} \quad J^*_{mj} = \sum_{i=1}^{n_w} \mathbb{1}\{U_{ji} \geq U_{i,(1)}(M_i(\mu^*))\}$$

Similarly, define the number of firm that worker i would accept and the number of workers that firm j would accept:

$$L^*_{wi} = \sum_{j=1}^{n_m} \mathbb{1}\{U_{ji} \geq U_{i,(1)}(M_i(\mu^*))\} \quad \text{and} \quad L^*_{mj} = \sum_{i=1}^{n_w} \mathbb{1}\{V_{ji} \geq V_{j,(q)}(W_j(\mu^*))\}$$

To characterize the limit of the conditional matching probabilities, we need to know at which rate these objects grow. Menzel (2015) showed that for $q = 1$, we can bound each of these by quantities that grow at a rate $\sqrt{n}$. I show that this extends to any $q > 1$ by proving the following:

Lemma 1 *Under Assumptions 1 and 2 and for any stable matching μ^*, we have for any finite $q \geq 1$:*

$$n^{1/2} \frac{\exp(-\bar{V} + \gamma_m)}{1 + \exp(\bar{U} + \bar{V} + \gamma_w)} \leq J^*_{wi} \leq n^{1/2} \exp(\bar{V} + \gamma_m)$$

$$n^{1/2} \frac{\exp(-\bar{U} + \gamma_w)}{1 + \exp(\bar{U} + \bar{V} + \gamma_m)} \leq J^*_{mj} \leq n^{1/2} \exp(\bar{U} + \gamma_w)$$

$$n^{1/2} \frac{\exp(-\bar{U} + \gamma_m)}{1 + \exp(\bar{U} + \bar{V} + \gamma_m)} \leq L^*_{wi} \leq n^{1/2} \exp(\bar{U} + \gamma_m)$$

$$n^{1/2}\frac{\exp(-\bar{V}+\gamma_w)}{1+\exp(\bar{U}+\bar{V}+\gamma_w)} \leq L^*_{mj} \leq n^{1/2}\exp(\bar{V}+\gamma_w)$$

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

PROOF: I rely on two important observations:

(a). We can bound, for any $q \geq 1$, $V_{j,(q)}(W_j(\mu^*))$ from above by $V_{j,(1)}(W_j(\mu^*))$ and similarly $U_{i,(q)}(M_i(\mu^*))$ by $U_{i,(1)}(M_i(\mu^*))$ for all $i = 1, ..., n_w$ and $j = 1, ..., n_m$.

(b). As in Menzel (2015), we can define exogenous sets $\bar{W}_j = \{i : U_{ij} \geq U_{i0}\}$ and $\bar{M}_i = \{j : V_{ij} \geq V_{0j}\}$ such that $W_j(\mu^*) \subset \bar{W}_j$ and $M_i(\mu^*) \subset \bar{M}_i$ as well as $W_j^\circ = \{i : U_{ij} \geq U_{i,(1)}(\bar{M}_i)\}$ and $M_{i,(q)}^\circ = \{j : V_{ij} \geq V_{j,(q)}(\bar{W}_j)\}$ such that $W_j^\circ \subset W_j(\mu^*)$ and $M_{i,(q)}^\circ \subset M_i(\mu^*)$.

A first important result is that (a) implies that for any $q > 1$, $M_{i,(1)}^\circ \subset M_{i,(q)}^\circ \subset M_i(\mu^*)$. From this, I construct the following bounds on J^*_{wi}:

$$J^\circ_{wi} = \sum_{j=1}^{n_m} \mathbb{1}\{j \in M_{i,(1)}^\circ\} \leq \sum_{j=1}^{n_m} \mathbb{1}\{j \in M_i(\mu^*)\} \leq \sum_{j=1}^{n_m} \mathbb{1}\{j \in \bar{M}_i\} = \bar{J}_{wi}$$

from there, using Proposition 2, we can show that:

$$\mathbb{E}[\bar{J}_{wi}|x_i, z_1, ..., z_{n_m}] = \frac{1}{J}\sum_{j=1}^{n_m} \frac{\exp\{V(x_i, z_j)\}}{1+\frac{1}{J}\exp\{V(x_i, z_j)\}} + o(1) \leq \frac{n_m}{J}\exp\{\bar{U}\} + o(1)$$

which implies under Assumption 2 that:

$$\mathbb{E}[\bar{J}_{wi}] \leq n^{1/2}\exp\{\bar{V}+\gamma_m\} + o(1)$$

Following the same steps as Menzel (2015) we can then show that the variance of $\bar{J}_{wi}$ converges to zero which implies that:

$$n^{-1/2}(\bar{J}_{wi} - \mathbb{E}[\bar{J}_{wi}]) \to 0$$

We have thus established that $J^*_{wi} \leq n^{1/2}\exp\{\bar{V}+\gamma_m\}$ with probability approaching 1 as $n \to \infty$. Following the same steps, we can show symmetrically that:

$$J^*_{mj} \leq n^{1/2}\exp\{\bar{U}+\gamma_w\}$$

$$L^*_{wi} \leq n^{1/2} \exp\{\bar{V} + \gamma_m\}$$

$$L^*_{mj} \leq n^{1/2} \exp\{\bar{U} + \gamma_w\}$$

with probability approaching 1 as $n \to \infty$. We now consider the lower bound J°_{wi}. We can again use Proposition 2 to show that:

$$\begin{aligned}\mathbb{E}[J^\circ_{wi}|(x_l)_{l\in\bar{W}_j}, z_1, ..., z_{n_m}] &= \frac{1}{J}\sum_{j=1}^{n_m} \frac{\exp\{V(x_i, z_j)\}}{1 + \frac{1}{J}\sum_{l\in\bar{W}_j}\exp\{V(x_l, z_j)\}} + o(1) \\ &\geq \frac{n_m}{J}\frac{\exp\{-\bar{V}\}}{1 + \frac{\bar{J}_{mj}}{J}\exp\{\bar{V}\}} + o(1)\end{aligned}$$

Using the higher bound for J^*_{mj} derived just above and Jensen's inequality, we can finally show that:

$$\mathbb{E}[J^\circ_{wi}] \geq n^{1/2}\frac{\exp\{-\bar{V} + \gamma_m\}}{1 + \exp\{\bar{V} + \bar{U} + \gamma_w\}} + o(1)$$

Following Menzel (2015) we can then also show that the variance of J°_{wi} converges to zero which implies that:

$$n^{-1/2}(J^\circ_{wi} - \mathbb{E}[J^\circ_{wi}]) \to 0$$

This establishes that $J^*_{wi} \geq n^{1/2}\frac{\exp\{-\bar{V}+\gamma_m\}}{1+\exp\{\bar{V}+\bar{U}+\gamma_w\}}$ with probability approaching 1 as $n \to \infty$. Following the same steps, we can show that symmetrically, we have:

$$J^*_{mj} \geq n^{1/2}\frac{\exp\{-\bar{U} + \gamma_w\}}{1 + \exp\{\bar{V} + \bar{U} + \gamma_m\}}$$

$$L^*_{wi} \geq n^{1/2}\frac{\exp\{-\bar{U} + \gamma_m\}}{1 + \exp\{\bar{V} + \bar{U} + \gamma_m\}}$$

$$L^*_{mj} \geq n^{1/2}\frac{\exp\{-\bar{V} + \gamma_w\}}{1 + \exp\{\bar{V} + \bar{U} + \gamma_w\}}$$

with probability approaching 1 as $n \to \infty$. This concludes the proof of Lemma 1.

A.3.2 Exogeneity of Feasible Choice Sets

We now need to show that as $n \to \infty$, the dependence between agents taste shocks and opportunity sets vanishes. Again, a proof exists for $q = 1$ in Menzel (2015) but I show that this result extends to $q > 1$.

For the rest of the proof, I define the following set of indicator functions $E^*_{ij} = \mathbb{1}\{i \in W_j(\mu^*)\}$ and $D^*_{ij} = \mathbb{1}\{j \in M_i(\mu^*)\}$ for all workers $i = 1, ..., n_w$ and firms $j = 1, ..., n_m$. The first result to establish is that the probability that changing one availability indicator affects another agents' opportunity set converges to zero as $n \to \infty$ for any $q \geq 1$. I first prove the following result:

Lemma 2 *Suppose Assumption 1 and 2 hold and suppose we change one availability indicator E^*_{ij} exogenously to $\tilde{E}_{ij} = 1 - E^*_{ij}$ and then iterate the deferred acceptance algorithm from this point until convergence. Denote the resulting availability indicators $\{\tilde{E}_{lk}, \tilde{D}_{lk} : l = 1, ..., n_w, k = 1, ..., n_m\}$. We have for any $q \geq 1$ and any worker l and firm k:*

(i). $\mathbb{P}(\tilde{D}_l \neq D^*_l | D^*_l, D^*_{ij} = 0) = \mathbb{P}(\tilde{E}_k \neq E^*_k | E^*_l, D^*_{ij} = 0) = 0$

(ii). There exist constants $\bar{a} < \infty$ and $0 < \lambda < 1$ such that:

$$\mathbb{P}(\tilde{D}_l \neq D^*_l | D^*_l, D^*_{ij} = 1) \leq n^{-1/2} \frac{\bar{a}}{1 - \lambda}$$

$$\mathbb{P}(\tilde{E}_k \neq E^*_k | E^*_l, E^*_{ij} = 1) \leq n^{-1/2} \frac{\bar{a}}{1 - \lambda}$$

The same result holds for an exogenous change of D_{ij} to $\tilde{D}_{ij} = 1 - D_{ij}$.

Proof: Suppose we change E^*_{ji} exogenously to $\tilde{E}_{ji} = 1 - E_{ji}$ and that we iterate the deferred acceptance algorithm from this stage. This will only trigger a chain of rematches if this affects the indirect utility of either i or j. Suppose $D^*_{ij} = 0$ and that $E^*_{ij} = 0$ meaning that firm j is not feasible to worker i and vice versa. Suppose now that $\tilde{E}_{ji} = 1 - E^*_{ij} = 1$, meaning that suddenly worker i's preference for firm j increase such that worker i becomes feasible for firm j. This will not affect the indirect utility of firm j nor worker i given that firm j is not feasible to worker i. This change will thus not trigger a chain of rematches. A similar argument can be used in the case where E^*_{ij} changes from 1 to $\tilde{E}_{ji} = 1 - E^*_{ij} = 0$. This establishes part (i) of Lemma 2 and does not depend on the value of q.

Now suppose that $D^*_{ij} = 1$ such that if $\tilde{E}_{ji} = 1 - E^*_{ij} = 1$, now firm j and worker i will want to rematch together or if $\tilde{E}_{ji} = 1 - E^*_{ij} = 0$ firm j and worker i will break their current match. This will trigger a chain of rematches than can potentially cycle back to worker i or firm j's opportunity set. I start by showing that, for any $q > 1$ at each step s of these subsequent rematches, there is at most one indicator in the vector $D_l^{(s)}$ corresponding to a firm k with $E_{lk}^{(s)} = 1$ that will change. The idea of the proof is the following: suppose that a given worker l matched to firm k in step $(s-1)$ becomes unavailable to firm k in step s. This firm will then replace this worker by the next q^{th} ranked feasible applicant, which will only change the availability indicator of this firm to this newly hired worker. On the other hand, if a given worker becomes available to a firm while this firm prefers this worker to one of its top q matched employees, then it will replace the q^{th} ranked worker by this new employee, making this firm unavailable to the kicked out employee. In both cases, this will only change at most one availability indicator among the workers who are willing to match with this firm. Note that at each of these steps, there is a chance that the chain is terminated if the next q^{th} ranked feasible worker is the outside option. A similar argument can be used symmetrically from the workers perspective by replacing q by 1.

The rest of the proof now consists in bounding the probability that the chain is terminated by either (a) firm k or worker l preferring the outside option to any other option in their opportunity set or (b) a change in availability indicators of worker k D_k. I define μ^s the state of the match in iteration s of the deferred acceptance algorithm following an exogenous change of E_{ji} to $\tilde{E}_{ji} = 1 - E_{ji}$. The first step bounds the probability that the chain is terminated by the outside option at stage s.

I start from the following observation: given that $\mathbb{P}(V_{lk} > V_{k,(q)}(W_k(\mu^s))|x_l, z_k) \geq \mathbb{P}(V_{lk} > V_{k,(1)}(W_k(\mu^s))|x_l, z_k)$ and that $W^\circ_{k,(1)} \subset W^*_k \subset \bar{W}_k$, we have from Proposition 2 and Lemma 1 that for any firm k and worker l:

$$\begin{aligned}\mathbb{P}(V_{lk} > V_{k,(q)}(W_k(\mu^s))|x_l, z_k) &\geq \mathbb{P}(V_{lk} > V_{k,(1)}(\bar{W}_k)|x_l, z_k) \\ &= n^{-1/2}\frac{\exp(V(z_k, x_l))}{1 + \frac{1}{J}\sum_{i\in\bar{W}_k}\exp(V(z_k, x_i))} + o(1) \\ &\geq n^{-1/2}\frac{\exp(V(z_k, x_l))}{1 + \exp(\bar{U} + \bar{V} + \gamma_w)} + o(1)\end{aligned}$$

This implies that, conditional on D_i^* and as n approaches infinity:

$$\mathbb{P}(V_{0k} > V_{k,(q)} | D_i^*, x_i, z_k) \geq \frac{1}{1 + \exp(\bar{U} + \bar{V} + \gamma_w)} =: p_s$$

Following now the same steps as Menzel (2015), we have, by Bayes law that:

$$\mathbb{P}(V_{0k} > V_{k,(q)} | D_l^*, \tilde{D}_{lk}^{(s)} = 1, x_l, z_k) \geq \frac{\underline{L} p_s}{\bar{L}(1 - p_s) + \underline{L} p_s}$$

where $\bar{L}$ and $\underline{L}$ are respectively the upper and lower bounds on L_{mj}^* taken from Lemma 1. From there, we finally get that:

$$1 - \mathbb{P}(V_{0k} > V_{k,(q)} | D_l^*, \tilde{D}_{lk}^{(s)} = 1, x_l, z_k) \leq \frac{\bar{L} \exp(\bar{U} + \bar{V} + \gamma_w)}{\bar{L} \exp(\bar{U} + \bar{V} + \gamma_w) + \underline{L}} =: \lambda < 1$$

This essentially means that the probability that the chain is not terminated at stage s is bounded away from 1.

Now we bound the probability that the chain leads to a change in D_l at stage s. We can thus bound the following probability using Proposition 2 and Lemma 1:

$$\begin{aligned}
&\mathbb{P}(V_{lk} > V_{k,(q)}(W_k(\mu^s)) | x_l, z_k) \\
&\quad \leq \mathbb{P}(V_{lk} > V_{k,(q)}(W_{k,(1)}^\circ) | x_l, z_k) \\
&\quad = n^{-1/2} \exp(V(z_k, x_l)) \left[1 - \left(\frac{\frac{1}{J} \sum_{i \in W_{k,(1)}^\circ} \exp(V(z_k, x_i))}{1 + \frac{1}{J} \sum_{i \in W_{k,(1)}^\circ} \exp(V(z_k, x_i))} \right)^q \right] + o(1) \\
&\quad \leq n^{-1/2} \exp(V(z_k, x_l)) \left[1 - \left(\frac{\frac{J_{mk}^\circ}{J} \exp(-\bar{V})}{1 + \frac{J_{mk}^\circ}{J} \exp(-\bar{V})} \right)^q \right] + o(1) \\
&\quad \leq n^{-1/2} \exp(V(z_k, x_l)) \left[1 - \left(\frac{\exp(-\bar{V} - \bar{U} + \gamma_w)}{1 + \exp(-\bar{V} - \bar{U} + \gamma_w) + \exp(\bar{U} + \bar{V} + \gamma_m)} \right)^q \right] + o(1) \\
&\quad \leq n^{-1/2} \exp(\bar{V}) + o(1)
\end{aligned}$$

This implies that for n sufficiently large, we have:

$$\begin{aligned}&\mathbb{P}(\tilde{D}_l^{(s)} \neq D_l^* | D_l^*, \tilde{D}_{lk}^{(s)} = 1, x_l, z_k) \\ &\leq \frac{n^{-1/2}\exp(\bar{V})\bar{L}}{n^{-1/2}\exp(\bar{V})\bar{L} + \underline{L}} \leq n^{-1/2}\exp(\bar{V})\frac{\bar{L}}{\underline{L}} = n^{-1/2}\bar{a}\end{aligned}$$

Using the law of total probability, we can thus bound as $n \to \infty$ the conditional probability that $\tilde{D}_l \neq D_l^*$:

$$\mathbb{P}(\tilde{D}_l \neq D_l^* | D_l^*) \leq \sum_{s=1}^{\infty} \lambda^s n^{-1/2}\bar{a} \leq \frac{n^{-1/2}\bar{a}}{1-\lambda}$$

which proves part (b) of Lemma 2.

From there, I state the main result that the dependence between taste shocks and agents' opportunity sets vanishes as $n \to \infty$ for any $q \geq 1$. I first define the joint distribution of $\eta_i = (\eta_{i1}, ..., \eta_{in_m})'$, $\epsilon_j = (\epsilon_{1j}, ..., \epsilon_{n_w j})'$ and the availability indicators D_i^W, E_j^W, D_i^M, E_j^M corresponding to the worker-optimal and the firm-optimal stable matches. Note that I consider these two specific matches since the worker-optimal and firm-optimal stable matches are defined with probability 1 conditional on the realization of the taste shocks η_i and ϵ_j. Indeed, the distribution of availability indicators arising from an arbitrary stable match D_i^* would not be well defined. I also define: $D_{i,-j}^W = (D_{i1}^W, ..., D_{i(j-1)}^W, D_{i(j+1)}^W, ..., D_{in_m}^W)$ and $E_{-i,j} = (E_{1j}^W, ..., E_{(i-1)j}^W, E_{(i+1)j}^W, ..., E_{n_w j}^W)$ with analogous notations for the firm optimal match. I then define the conditional c.d.f.s:

$$G_{\eta|D}^W(\eta|\boldsymbol{d}) = \mathbb{P}(\eta_i \leq \eta | D_i^W = \boldsymbol{d}), \quad \boldsymbol{d} \in \{0,1\}^{n_m}$$

$$G_{\eta,\epsilon|D,E}^W(\eta, \epsilon|\boldsymbol{d}, \boldsymbol{e}) = \mathbb{P}(\eta_i \leq \eta, \epsilon_j \leq \epsilon | D_{i,-j}^W = \boldsymbol{d}, E_{-i,j}^W = \boldsymbol{e}), \quad \boldsymbol{d} \in \{0,1\}^{n_m - 1}, \boldsymbol{e} \in \{0,1\}^{n_w - 1}$$

with analogous definitions for the firm-optimal stable match and associated p.d.f.s $g_{\eta|D}^W$ and $g_{\eta,\epsilon|D,E}^W$. The main result is the following:

Lemma 3 *Under Assumption 1 and 2, we have:*

(i). $g^W_{\eta|D}$ and $g^M_{\eta|D}$ satisfy:

$$\lim_n \Big| \frac{g^W_{\eta|D}(\eta|D^W_i)}{g_\eta(\eta)} - 1 \Big| = \lim_n \Big| \frac{g^M_{\eta|D}(\eta|D^M_i)}{g_\eta(\eta)} - 1 \Big| = 1$$

(ii). $g^W_{\eta,\epsilon|D,E}$ and $g^M_{\eta,\epsilon|D,E}$ satisfy:

$$\lim_n \Big| \frac{g^W_{\eta|D}(\eta,\epsilon|D^W_{i,-j},E^W_{-i,j})}{g_{\eta,\epsilon}(\eta,\epsilon)} - 1 \Big| = \lim_n \Big| \frac{g^M_{\eta|D}(\eta,\epsilon|D^M_{i,-j},E^M_{-i,j})}{g_{\eta,\epsilon}(\eta,\epsilon)} - 1 \Big| = 1$$

The same results holds for the firm side of the market.

PROOF: Let $g^W_{\eta,D}$ be the joint p.d.f. of taste shocks and availability indicators under the worker optimal stable match. We can rewrite, by definition of a conditional density:

$$\frac{g^W_{\eta|D}(\eta|D^W_i)}{g_\eta(\eta)} = \frac{g^W_{\eta,D}(\eta, D^W_i)}{g_\eta(\eta)P(D^W_i)} = \frac{P(D^W_i|\eta_i = \eta)g_\eta(\eta)}{g_\eta(\eta)P(D^W_i)} = \frac{P(D^W_i|\eta_i = \eta)}{P(D^W_i)}$$

I then follow similar steps as in Menzel (2015) to show that:

$$\Big| \frac{P(D^W_i|\eta_i = \eta)}{P(D^W_i)} - 1 \Big| \leq \sup_{\eta_1,\eta_2} \Big| \frac{P(D^W_i|\eta_i = \eta_1)}{P(D^W_i|\eta_i = \eta_2)} - 1 \Big|$$

such that I only need to bound the probability that shifting η_i from η_1 to η_2 changes worker i's opportunity set. This insight does not depend on q. We know from Lemma 2 that changing an availability indicator will trigger a chain of rematches that could change worker i's opportunity set with probability less than $\frac{n^{-1/2}\bar{a}}{1-\lambda}$ as n approaches infinity. Here, we can show that shifting agent i's taste shocks would trigger at most two chains of rematches. Indeed, if the shift in taste shocks makes agent i prefers firm l with $D_{il} = 1$ instead of her current employer firm j, this changes both E_{ij} from 1 to 0 and E_{il} from 0 to 1. Thus, this would trigger two chains of rematches where both firm j and the worker which was displaced from firm l by worker i would need to find a new match. We can thus conclude that:

$$\frac{P(D^W_i|\eta_i = \eta_1)}{P(D^W_i|\eta_i = \eta_2)} - 1 \leq 2\frac{n^{-1/2}\bar{a}}{1-\lambda}$$

as $n \to \infty$ which can be shown to hold also in absolute value. As the right hand side converges to 0 as $n \to \infty$, this proves the first part of claim (i). Now consider the symmetrical case where we would shift firm j's taste shocks. Following a similar argument, we can see that this would create at most $2q$ chains of rematches. Indeed, assuming that such a shift in firm j's taste shocks would make it want to replace all of its q employees, this implies that the q workers which were let go along with the (potentially) q firms which lost one of their employees would need to find a new match. This implies that:

$$\frac{P(E_j^W|\epsilon_j = \epsilon_1)}{P(E_j^W|\epsilon_j = \epsilon_2)} - 1 \leq qn^{-1/2}\frac{2\bar{a}}{1-\lambda}$$

as $n \to \infty$ which can be shown to hold also in absolute value. As the right hand side converges to 0 as $n \to \infty$, this proves the first part of claim (i).

For part (ii), note that the argument can be extended in a similar way. If you change both firm j and worker i's taste shocks this can trigger at most $2(q+1)$ chains of rematches such that we can bound the probability of a shift in opportunity sets by $(q+1)n^{-1/2}\frac{2\bar{a}}{1-\lambda}$ which can be made arbitrarily close to 0 as n approaches infinity.

A.3.3 Bounds for Inclusive Values

Since I have established exogeneity of opportunity sets under the firm-optimal and worker-optimal stable matches, the rest of the analysis focuses on characterizing the limit of inclusive values that arise under these extremal matchings. As in Menzel (2015), I show that both converge to a unique limit, implying that inclusive values arising from any stable matching also converge towards this limit.

I define $I_{wi}^W = I_{wi}(\mu^W)$ and $I_{mj}^W = I_{mj}(\mu^W)$ the inclusive values that arise from the worker-optimal stable match. Similarly, I define I_{wi}^M and I_{mj}^M as the inclusive values that arise from the firm-optimal stable match such that for any stable match μ^*, we have $I_{wi}^W \geq I_{wi}(\mu^*) \geq I_{wi}^M$ and $I_{mj}^W \leq I_{wi}(\mu^*) \leq I_{mj}^M$. I state the following result:

Lemma 4 *Under Assumption 1 and 2:*

(i). For all $i = 1, ..., n_w$ and $j = 1, ..., n_m$:

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

where the analogous result holds for the worker-optimal stable match with the side of inequalities reversed.

(ii). If the weight functions $\omega(x, z) \geq 0$ are bounded and form a Glivenko-Cantelli class in x, then

$$\sup_{x \in \mathcal{X}} \frac{1}{n} \sum_{j=1}^{n_m} \omega(x, z_j)(I_{mj}^M - \hat{\Gamma}_m^M(z_j)) \leq o_p(1)$$

and

$$\inf_{z \in \mathcal{Z}} \frac{1}{n} \sum_{i=1}^{n_w} \omega(x_i, z)(I_{wi}^M - \hat{\Gamma}_w^M(x_i)) \geq o_p(1)$$

The analogous conclusion holds for the worker-optimal stable match where the sign of the inequalities is reversed and if $\omega(x, z) \geq 0$ are bounded and form a Glivenko-Cantelli class in z.

Proof: I first show that we can bound conditional choice probabilities given an opportunity set arising from a stable match using the extremal matchings. I first define the conditional probability that worker i chooses firm j given the realization of opportunity set M^M arising from the firm-optimal stable match:

$$\Lambda_w^M(x, z, M^M) = \mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i^M) | M_i^M = M^M, x_i = x, z_j = z)$$

Similarly, I define the equivalent object from firm j's perspective:

$$\Lambda_m^M(x, z, W^M) = \mathbb{P}(V_{ij} \geq V_{j,(q)}(W_j^M) | W_j^M = W^M, x_i = x, z_j = z)$$

I also define the conditional choice probabilities under exogenous opportunity sets as:

$$\Lambda_w(x, z, M) = \mathbb{P}(U_{ij} \geq U_{i,(1)}(M) | x_i = x, z_j = z)$$

$$\Lambda_m(x, z, W) = \mathbb{P}(V_{ij} \geq V_{i,(q)}(W) | x_i = x, z_j = z)$$

As there are several stable matches such that $M_i^* = M_i^M$ and $W_j^* = W_j^M$ we can show that:

$$J\Lambda_w^M(x, z, M_i^M) \leq J\Lambda_w(x, z, M_i^M) + o_p(1)$$

$$J\Lambda_m^M(x, z, W_j^M) \geq J\Lambda_m(x, z, W_j^M) + o_p(1)$$

Similarly, we have:

$$J\Lambda_w^W(x, z, M_i^W) \geq J\Lambda_w(x, z, M_i^W) + o_p(1)$$

$$J\Lambda_m^W(x, z, W_j^W) \leq J\Lambda_m(x, z, W_j^W) + o_p(1)$$

Using Proposition 2, we can then show that for $i = 1, ..., n_w$, $l_1 = 1, ..., n_m$ and $l_2 \neq l_1$:

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

and

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

Therefore, since under Assumption 1, we know that $\exp(U(x_i, z_j))$ is bounded, we can thus conclude that:

$$\text{Var}\left(\frac{1}{n}\sum_{k=1}^{n_m} \exp\{U(x_i, z_k)\} J(D_{ik}^M - \Lambda_m^M(x_i, z_k, I_{mk}^M))\right) \to 0$$

which implies that:

$$\frac{1}{n}\sum_{k=1}^{n_m} \exp\{U(x_i, z_k)\} J(D_{ik}^M - \Lambda_m^M(x_i, z_k, I_{mk}^M)) = o_p(1)$$

Given that from Proposition 2:

$$J\Lambda_m^M(x, z, W_j^M) \geq \exp\{V(x, z)\}\left[1 - \left(\frac{I_{mj}^M}{1 + I_{mj}^M}\right)^q\right] + o_p(1)$$

This implies that:

$$\frac{1}{n}\sum_{k=1}^{n_m}\exp\{U(x_i,z_k)\}\left(JD_{ik}^M-\exp\{V(x_i,z_k)\}\left[1-\left(\frac{I_{mk}^M}{1+I_{mk}^M}\right)^q\right]\right)\geq o_p(1)$$

which proves the first claim of part (i) of Lemma 4. From firm j's perspective, I show using the same arguments that:

$$\frac{1}{n}\sum_{l=1}^{n_w}\exp\{V(x_l,z_j)\}J(E_{lj}^M-\Lambda_w^M(x_l,z_j,I_{wl}^M))=o_p(1)$$

which implies, using Proposition 2:

$$\frac{1}{n}\sum_{l=1}^{n_w}\exp\{V(x_l,z_j)\}\left(JE_{lj}^M-\frac{\exp\{U(x_l,z_j)\}}{1+I_{wl}^M}\right)\geq o_p(1)$$

This prove part (i) of Lemma 4. Note that analogous arguments can be used to bound inclusive values arising from the worker-optimal stable match.

Part (ii) follows from part (i) of the Lemma and the boundedness condition on ω which implies pointwise convergence. The Glivenko-Cantelli condition on ω then implies uniform convergence. This concludes the proof of Lemma 4.

The next step consists in establishing uniform convergence with respect to $\Gamma_w\in\mathcal{T}_w$ and $\Gamma_m\in\mathcal{T}_m$ of the fixed point mappings $\hat{\Psi}_w$ and $\hat{\Psi}_m$ to their population counterparts. I define:

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

where ψ_w is defined as:

$$\psi_w(z_j,x;\Gamma_m)=\exp\{U(x,z_j)+V(x,z_j)\}\left[1-\left(\frac{\Gamma_m(z_j)}{1+\Gamma_m(z_j)}\right)^q\right]$$

Similarly, I define:

$$\hat{\Psi}_m[\Gamma_w](z)=\frac{1}{n}\sum_{i=1}^{n_w}\psi_m(z,x_i;\Gamma_w)$$

where ψ_m is defined as:

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

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

Lemma 5 *Under Assumption 1:*
(i). The classes of functions $\mathcal{F}_w$ and $\mathcal{F}_w$ are Glivenko-Cantelli.
(ii). As $n \to \infty$:

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

uniformly in $\Gamma_w \in \mathcal{T}_w$, $\Gamma_m \in \mathcal{T}_m$ and $(x, z) \in \mathcal{X} \times \mathcal{Z}$.

Proof: Under Assumption 1, $\exp\{U(x, z) + V(x, z)\}$ is Lipschitz in x and z such that this class of functions is Glivenko-Cantelli. Γ_m and Γ_w are bounded and have bounded $p \geq 1$ derivatives which makes the class of functions $\mathcal{F}_w \cup \mathcal{F}_m$ Glivenko-Cantelli. Finally, note that the transformation $\psi_m(g, h) = \frac{g}{1+h}$ is bounded and continuous since h and g are bounded and continuous and $h \geq 0$. Similarly, the transformation $\psi_w(g, h) = g\left[1 - \left(\frac{h}{1+h}\right)^q\right]$ is also bounded and continuous for any $q \geq 1$. Theorem 3 in van der Vaart and Wellner (2000) implies claim (i) of Lemma 5. Part (ii) of Lemma 5 is a direct implication of part (i).

A.3.4 Proof of Theorem 3.1 (iii)

I finally turn to the proof of part (iii) of Theorem 1. I first apply Lemma 4 to show that for any $q \geq 1$:

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

Similarly, I show that:

$$\hat{\Gamma}_m^M(z) \leq \frac{1}{n}\sum_{i=1}^{n_w} \frac{\exp\{U(x_i,z)+V(x_i,z)\}}{1+\hat{\Gamma}_w^M(x_i)} + o_p(1)$$

Analogous bounds can be formed for the inclusive value functions of the worker-optimal stable match. We thus have that:

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

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

Given that $\hat{\Psi}_w[\Gamma_m]$ and $\hat{\Psi}_m[\Gamma_w]$ are nonincreasing and Lipschitz continuous in Γ_m and Γ_w, we have:

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

Thus for any pair of functions (Γ_w^*, Γ_m^*) solving the fixed point problem:

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

we thus have:

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

However, we know that the mapping $\hat{\Psi}$ is a contraction in logs, which means that it has a unique fixed point (Γ_w^*, Γ_m^*). We also know, by definition, that:

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

which implies that:

$$\Gamma_w^* + o_p(1) \geq \hat{\Gamma}_w^W \geq \hat{\Gamma}_w^M \geq \Gamma_w^* + o_p(1)$$

$$\Gamma_m^* + o_p(1) \leq \hat{\Gamma}_m^W \leq \hat{\Gamma}_m^M \leq \Gamma_m^* + o_p(1)$$

which in turn implies that:

$$\hat{\Gamma}^M_w = \Gamma^*_w + o_p(1) \quad \text{and} \quad \hat{\Gamma}^M_m = \Gamma^*_m + o_p(1)$$

$$\hat{\Gamma}^W_w = \Gamma^*_w + o_p(1) \quad \text{and} \quad \hat{\Gamma}^W_m = \Gamma^*_m + o_p(1)$$

Combining this with Lemma 3, this gives us for all $i = 1, ..., n_w$ and all $j = 1, ..., n_m$:

$$I^M_{wi} = \Gamma^*_w + o_p(1) \quad \text{and} \quad I^M_{mj} = \Gamma^*_m + o_p(1)$$

$$I^W_{wi} = \Gamma^*_w + o_p(1) \quad \text{and} \quad I^M_{mj} = \Gamma^*_m + o_p(1)$$

Note that given that inclusive value functions that would arise under any stable match μ^* defined as I^*_{wi} and I^*_{mj} are such that $I^M_{wi} \leq I^*_{wi} \leq I^W_{wi}$ and $I^M_{mj} \geq I^*_{mj} \geq I^W_{mj}$ the equality written above holds also for any I^*_{wi} and I^*_{mj}.

I have shown that inclusive values can be approximated by the solution of the finite sample fixed point problem. Lemma 5 finally implies that the solution of the finite sample fixed point problem converges towards the solution of its population equivalent. This proves Theorem 1.(iii).

A.4 Proof of Proposition 3

Assume that firm j is matched with a group of $k = 2$ workers and that we want to characterize the limit of the following CCP:

$$\mathbb{P}(\mu_m(j) = \{i, l\} \cup \{0\}^{q-2} | x_i, x_l, z_j)$$

We can rewrite it as follows:

$$\begin{aligned}
&\mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^*)), U_{lj} \geq U_{l,(1)}(M_l(\mu^*)), V_{ij} > V_{lj} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*)) | x_i, x_l, z_j) \\
&+ \mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^*)), U_{lj} \geq U_{l,(1)}(M_l(\mu^*)), V_{lj} > V_{ij} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*)) | x_i, x_l, z_j) \\
=&\mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^*)) | x_i, z_j) \times \mathbb{P}(U_{lj} \geq U_{i,(1)}(M_i(\mu^*)) | x_l, z_j) \\
&\times \left[\mathbb{P}(V_{ij} > V_{lj} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*)) | x_i, x_l, z_j) + \mathbb{P}(V_{lj} > V_{ij} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*)) | x_i, x_l, z_j)\right]
\end{aligned}$$

We can then decompose the rank ordered CCPs as follows:

$$\begin{aligned}
&\mathbb{P}(V_{ij} > V_{lj} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*))|x_i, x_l, z_j) \\
=&\mathbb{P}(V_{ij} \geq V_{j,(1)}(W_j(\mu^*))|x_i, x_l, z_j) \times \mathbb{P}(V_{lj} \geq V_{j,(1)}(W_j(\mu^*) \setminus i)|x_i, x_l, z_j) \\
&\times \mathbb{P}(V_{0j} \geq V_{j,(1)}(W_j(\mu^*) \setminus \{i, l\})|x_i, x_l, z_j)
\end{aligned}$$

In the limit, removing one arbitrary alternative from opportunity sets does not affect inclusive values:

$$\begin{aligned}
n^{-1/2} \sum_{i \in W_j(\mu^*) \setminus \{l\}} \exp\{V(x_i, z_j)\} &= n^{-1/2} \sum_{i \in W_j(\mu^*)} \exp\{V(x_i, z_j)\} - n^{-1/2} \exp\{V(x_l, z_j)\} \\
&= n^{-1/2} \sum_{i \in W_j(\mu^*)} \exp\{V(x_i, z_j)\} + o_p(1) = I^*_{mj} + o_p(1)
\end{aligned}$$

We can thus conclude that:

$$\begin{aligned}
&\mathbb{P}(V_{ij} > V_{lj} > V_{0j} \geq V_{j,(3)}(W_j(\mu^*))|x_i, x_l, z_j) \\
=&\mathbb{P}(V_{ij} \geq V_{j,(1)}(W_j(\mu^*))|x_i, x_l, z_j) \times \mathbb{P}(V_{lj} \geq V_{j,(1)}(W_j(\mu^*))|x_i, x_l, z_j) \\
&\times \mathbb{P}(V_{0j} \geq V_{j,(1)}(W_j(\mu^*))|x_i, x_l, z_j) + o(1)
\end{aligned}$$

Which, using Proposition 2 and Theorem 1, implies that:

$$n^2 \mathbb{P}(\mu_m(j) = \{i, l\} \cup \{0\}^{q-2}|x_i, x_l, z_j) \longrightarrow \frac{2 \exp\{U(x_i, z_j) + U(x_l, z_j) + V(x_i, z_j) + V(x_l, z_j)\}}{(1 + \Gamma^*_w(x_i))(1 + \Gamma^*_w(x_l))(1 + \Gamma^*_w(z_j))^3}$$

where Γ^*_w and Γ^*_m are the solutions of the fixed point problem described in Equation 6. To extend to proof to any k, a similar argument applies, except that the number of rank ordered CCPs becomes $k!$. This proves part (i) of Proposition 3.

A similar argument can be used to prove part (ii).

A.5 Proof of Proposition 6

From Theorem 1 and Proposition 2, we know that for any $q \geq 1$ and for a given finite w:

$$n^{-1/2}\mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^*))|x_i = x, z_j = z, w_i = w) \to \frac{\exp\left(U(x,z)\right)}{1 + \Gamma^*_w(x,w)}$$

and

$$J\mathbb{P}(V_{ij} \geq V_{j,(1)}(W_j(\mu^*))|x_i = x, z_j = z, w_i = w) \to \frac{\exp\left(V(x,z) + g(w)\right)}{1 + \Gamma_m(z)}$$

where Γ^*_m and Γ^*_w solve the following fixed point problem:

$$\Gamma^*_w = \Psi_w[\Gamma^*_m] \quad \text{and} \quad \Gamma^*_m = \Psi_m[\Gamma^*_w] \tag{7}$$

where

$$\Psi_w[\Gamma_m](x,w) = \int \exp(U(x,s) + V(x,s) + g(w) + \gamma_m) \times \left[1 - \left(\frac{\Gamma_m(s)}{1 + \Gamma_m(s)}\right)^q\right] m(s)ds$$

$$\Psi_m[\Gamma_w](x) = \int\int \frac{\exp(U(s,z) + V(s,z) + g(t) + \gamma_w)}{1 + \Gamma_w(s,t)} w_x(s) w_w(t)ds$$

However, as w goes to infinity, we have that:[14]

$$\lim_{w \to \infty} \mathbb{P}(V_{ij} \geq V_{j,(q)}(W_j(\mu^*))|x_i = x, z_j = z, w_i = w) = 1$$

This implies that:

$$\lim_{w_i \to \infty} n^{-1/2} I_{wi} = \int \exp\{U(x,s) + \gamma_m\} m(s)ds$$

which in turn implies that:

$$\lim_{w \to \infty} n\mathbb{P}(U_{ij} \geq U_{i,(1)}(M_i(\mu^*))|x_i = x, z_j = z, w_i = w) = \frac{\exp\left(U(x,z)\right)}{1 + \int \exp\{U(x,s) + \gamma_m\} m(s)ds}$$

[14]The probability that an option is in workers' opportunity sets only goes to 1 when making the shifter w go to infinity. In He et al. (2021) and Agarwal and Somaini (2022) this is not the case as cutoffs are fixed and finite since the number of "colleges" or "products" is fixed. As both the number of firms and workers grow to infinity in our case, the cutoffs grow to infinity as the size of the market grows.

Similarly,

$$\lim_{w \to \infty} n\mathbb{P}(U_{i0} \geq U_{i,(1)}(M_i(\mu^*))|x_i = x, z_j = z, w_i = w) = \frac{1}{1 + \int \exp\{U(x,s) + \gamma_m\} m(s) ds}$$

Taking the log of these ratios separately identifies U from the joint surplus. Given that the joint surplus is identified for finite w, we can then recover $V + g$. V can be separately identified by evaluating $V + g$ at $\bar{w}$.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak, "The Welfare Effects of Coordinated Assignment: Evidence from the New York City High School Match," *American Economic Review*, 2017, *107* (12), 3635–89.

Agarwal, Nikhil, "An Empirical Model of the Medical Match," *American Economic Review*, 2015, *105* (7), 1939–78.

_ , "Policy Analysis in Matching Markets," *American Economic Association P&P*, 2017, *107* (5), 246–50.

_ **and Eric Budish**, "Market Design," in "Handbook of Industrial Organization," Vol. 5, Elsevier, 2021, pp. 1–79.

_ **and Paulo J Somaini**, "Demand Analysis under Latent Choice Constraints," *NBER Working Paper 29993*, 2022.

_ **and Paulo Somaini**, "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 2020, *12*, 471–501.

_ **and** _ , "Empirical Models of Non-Transferable Utility Matching," in "Online and Matching-Based Market Design" Forthcoming.

Azevedo, Eduardo M and Jacob D Leshno, "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, 2016, *124* (5), 1235–1268.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

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

Diamond, William and Nikhil Agarwal, "Latent Indices in Assortative Matching Models," *Quantitative Economics*, 2017, *8* (3), 685–728.

Fox, Jeremy T, "Identification in Matching Games," *Quantitative Economics*, 2010, *1* (2), 203–254.

Galichon, Alfred and Bernard Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *Review of Economic Studies*, Forthcoming.

Gualdani, Cristina and Shruti Sinha, "Partial Identification in Matching Models for the Marriage Market," *arXiv:1902.05610*, 2019.

He, YingHua, Shruti Sinha, and Xiaoting Sun, "Identification and Estimation in Many-to-one Two-sided Matching Without Transfers," *arXiv:2104.02009*, 2021.

Judd, Kenneth L, *Numerical Methods in Economics*, MIT press, 1998.

Kapor, Adam J, Christopher A Neilson, and Seth D Zimmerman, "Heterogeneous Beliefs and School Choice Mechanisms," *American Economic Review*, 2020, *110* (5), 1274–1315.

Menzel, Konrad, "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 2015, *83* (3), 897–941.

_ , "Bootstrap With Cluster-Dependence in Two or More Dimensions," *Econometrica*, 2021, *89* (5), 2143–2188.

Resnick, Sidney I, *Extreme Values, Regular Variation, and Point Processes*, Vol. 4, New York: Springer, 1987.

Roth, Alvin E, "Marketplaces, Markets, and Market Design," *American Economic Review*, 2018, *108* (7), 1609–58.

_ **and Marilda Sotomayor**, "Two-Sided Matching," *Handbook of Game Theory with Economic Applications*, 1992, *1*, 485–541.