Identification in Matching Models with Search Friction *

Lixiong Li†

October 24, 2015

Abstract

I investigate a model of one-to-one matching with transferable utilities, where the matching process is subject to time-consuming search frictions. I assume agents have unobserved (to economists) characteristics, which affect the matching surplus along with matching specific random shocks under a separability assumption. I show the matching surplus can be non-parametrically identified with data on matching patterns and distributions on unmatched durations across agents, given any known distribution on unobserved characteristics. Different from existing literature, my identification strategy does not hinge on data on payoffs and panel data with long time series. As in frictionless matching models, I show any interior matching patterns can be rationalized by the model under some parameters. For one type of corner solution, only set identification is attained and a sharp bound has been derived.

*I am grateful to Marc Henry for his suggestion and support to carry this paper out. For helpful comments, I thank Martin Hackmann, Paul Grieco, Kala Krishna, Shouyong Shi and Neil Wallace.

†Department of Economics, Pennsylvania State University, email: lixiong@psu.edu.

1 Introduction

Matching problems are pervasive. Some markets, like marriage market and labor market, can usually be characterized by a one-to-one matching model with transferable utility and time-consuming search friction. One challenge for empirical research in these markets, is that economists cannot observe all characteristics of agents which agents themselves can observe. In fact, as discussed in Wong (2003), some kinds of unobserved characteristics are often needed to get meaningful results from the data.

There have been recent advances in empirical labor search literature on search and matching models. In particular, Hagedorn et al. (2014) and Lamadon et al. (2013) have developed methods to identify matching surplus in a search-matching-bargain framework where types of agents are unobserved. Yet for two reasons, the full empirical content of such search-matching models remains to be exploited.

First, the identification in these papers hinges on a panel data with long time dimension. In their simulation examples, Hagedorn et al. (2014) assumes a panel with 10-20 years at monthly frequency, and Lamadon et al. (2013) assumes a panel with 40 years at quarterly frequency. The demand for a long panel has two consequences: (i) their methods might not work well when only a panel of limited time dimension is available; (ii) The stationary equilibrium assumption may be inappropriate when applied to data of decades long, due to possible long-term trend in model primitives.

Second, the payoff data (wage) is essential to the identification strategies developed in these papers. Although most matched-employer-employee database contains wage data, dependence on payoff data limits their methods to be applied to other markets, for example, marriage markets, where transfers between family members are rarely observable. Also, workers may value other factors in addition to wage in some cases (see, for example, Boyd et al. (2013)) so that worker's payoff could be unobserved even with the availability of wage data.

This paper establishes nonparametric identification of matching surplus in a search-matching-bargain framework as Shimer and Smith (2000), where agents have both observable and unobserved characteristics, and are subject to matching-specific random shocks. Given distributions of unobserved characteristics, the identification only relies on two kinds of data: (i) who matches with whom and (ii) distribution of unmatched duration across agents.

The main challenge here is the endogeneity of conditional distributions on unobserved characteristics among unmatched agents. Due to the selection bias, agents in the unmatched pool are, less "attractive" in general. Such discrepancy is essential, since it is the distribution in the unmatched pool that matters for the emergence of new matched pairs. Since we do not assume the availability of payoff data and a long panel data, we cannot identify the unobserved types of each agent directly as in Hagedorn et al. (2014) and Lamadon et al. (2013). As a result, such endogenous conditional distribution is, at the same time, unobserved, which poses new challenges to the identification problem.

To deal with this problem, we make two key assumptions: (i) observable and unobserved characteristics are separable in matching surplus, as assumed in Choo and Siow (2006) and Galichon and Salanié (2014); (ii) the unobserved heterogeneity is one-dimensional and, hence, can be totally ordered within each group of agents. Under these two assumptions, agent's hazard rate of getting matched is non-decreasing with unobserved types, so that we can identify the conditional distribution among unmatched agents constructively.

The reminder of the paper is organized as follows. Related Literature is discussed in Section 2. Section 3 describes the economic environment of search-matching-bargain framework, while Section 4 sets up the econometric problem, introduces observable and unobserved characteristics, and explicitly describes the data and assumptions needed. Section 5 studies a special case where identification can be easily established, and discuss the relation between frictional and frictionless matching. Section 6 exhibits our main results on the nonparametric identification of matching surplus. Finally, two extensions are discussed in Section 7. Some proofs and technique details are left in the Appendix.

2 Related Literature

In this section, we briefly discuss the closely related literature on econometrics of matching and search models. We mainly focus on papers with transferable utility matching models.

Our paper is related to the literature on static and frictionless matching models with transferable utilities. In particular, our identification method is related to Choo and Siow (2006) and Galichon and Salanié (2014). As in these two paper, we treat observable characteristics as discrete groups and recover matching surplus with (group-pairwise) matching patterns under separability assumption. Yet, Choo and Siow (2006) and Galichon and Salanié (2014) consider a static frictionless matching market and equilibrium can be characterized as social optimal solution, while we consider a matching model with search frictions in a continuous time setting where the solution concept is stationary equilibrium. As a result, we have very different identification conditions and treatments. There are, of course, other papers in these fields where identification methods can also be index based (Chiappori et al. (2012)) or rank-order based (Fox (2010), Fox and Bajari (2013)). Chiappori and Salanié (2015) serves a good survey in this literature.

Our paper is also related to recent works on the dynamic matching models. See, for example, Choo and Siow (2007), Choo (2015) and Bruze et al. (2015). In these papers, there are multiple time period, yet the matching market clears out frictionlessly in each period. Agents need to make trade-offs between matching today and matching in the future. Such trade-offs make sense as both global and individual states change over time. Though also studying a matching models with dynamics, we focus on markets with time-consuming search frictions. Dynamic matching models emphasis the impact of changes in economic environment like changes in sex-ratio or age-structure, yet search and matching models care about how frictions will shape the stationary matching patterns. Which model is more suitable mainly depends on the research questions.

Finally, this paper is closely related to the literature on matching models with search frictions, among which there are two types. The first type in this literature assumes (i) all characteristics of agents are observable, and (ii) identification results can be achieved with data on who matches with whom, matched and unmatched duration. Currently, most of these papers focus on the marriage market. For example, Jacquemet and Robin (2012) assumes the type of agents is their ability observed as their wages in the labor market and study the interaction between labor supply and marriage formation. Shin (2014) assumes ethnic groups as the sole characteristic of agents and study the reason behind racial homogamy. Wong (2003) and Goussé (2014) assume agent's types is an index based on education, wage income and physical attractiveness. Although it seems now standard in this literature to assume matching specific random shocks yet ignore unobserved heterogeneity among agents, there is no particular good reason to do so, as unobserved heterogeneity plays an essential role in most frictionless matching literature. In this paper, we include both matching specific shocks and agent's unobserved characteristics as the random components.

Contrary to the first type, the second type of empirical search and matching literature assumes (i) all characteristics of agents are unobserved, and (ii) identification results hinges on the availability of payoff data (wage) and a long panel data. Most of these works focus on the labor market. Eeckhout and Kircher (2011) shows the matching surplus cannot be identified with wage data in a model with explicit search cost, while Hagedorn et al. (2014) and Lamadon et al. (2013) show the matching surplus can be nonparametrically identified in a model with time-consuming search cost. As discussed in the introduction, we differ with these papers in less demanding data requirement and no need for payoff data.

3 Economic Matching Models with Search Fictions

Consider a bipartite matching model with search frictions as in Shimer and Smith (2000). Time is continuous. There are two kinds of infinitely lived agents in the market. Following the tradition of the literature, we call them men and women, though the implications of this paper are clearly not restricted to the marriage market.

Each man $i \in I$ has his characteristics $\tilde{x}_i \in \mathbb{R}^k$, and each woman $j \in J$ has her characteristics $\tilde{y}_j \in \mathbb{R}^k$, where $k \geq 1$, and I and J be the index set of men and women. Let $\mathbb{G}$ and $\mathbb{Q}$ denote the measure of men in I and women in J, and let $\mathcal{G}$ and $\mathcal{Q}$ denote the measure of their characteristics in $\mathbb{R}^k$, respectively. By definition, the measure of women with characteristics in a measurable set B is $\mathcal{Q}(B) = \mathbb{Q}(\{j \in J : \tilde{y}_j \in B\})$. The support of $\mathcal{G}$ and $\mathcal{Q}$ is a subset of $\mathbb{R}^k$.

As in Shimer and Smith (2000), it's without of loss of generality to assume agents with the same characteristics adopt the same strategy, and have the same expected present value of payoffs *ex ante*. The equilibrium under this assumption remains to be an equilibrium without this assumption. When there is no ambiguity, we also index man with his characteristics $\tilde{x}$ and woman with her characteristics $\tilde{y}$.

Each agent maximizes his/her expected present value of payoffs, discounted at interest rate r. Agents have two possible states: either matched or unmatched.

When unmatched, agents receive a constant utility flow, and attempt to match with others subject to search frictions. For man $\tilde{x}$ and woman $\tilde{y}$, the utility flow they receive when unmatched are denoted by $b(\tilde{x})$ and $c(\tilde{y})$. To form a matching, an agent must wait to have a chance to meet the opposite side and matching is formed if they both accept each other in a meeting. One agent can only meet one agent from the opposite sex at a time. Let u and v denote the total measure of unmatched men and women respectively. Assume the flow of meetings between men and women is given by a matching technology $m(u, v)$. Thus, the Poisson rate of a man meeting some woman is $m(u, v)/u$, and the rate of a woman meeting some man is $m(u, v)/v$. When man $\tilde{x}$ and woman $\tilde{y}$ meet, they learn their matching surplus $f(\tilde{x}, \tilde{y})$ instantly, and decide (i) whether to accept each other; (ii) how to split the matching surplus if both agree to form a matching. The acceptance decisions of man $\tilde{x}$ and woman $\tilde{y}$ are described by their acceptance set,

$$\mathcal{A}_{\tilde{x}} = \left\{\tilde{y} \in \mathbb{R}^k : \text{Woman } \tilde{y} \text{ will be accepted by man } \tilde{x}\right\}$$
$$\mathcal{A}_{\tilde{y}} = \left\{\tilde{x} \in \mathbb{R}^k : \text{Man } \tilde{x} \text{ will be accepted by woman } \tilde{y}\right\}$$

and a matching is formed if man $\tilde{x}$ and woman $\tilde{y}$ meet and $\tilde{x} \in \mathcal{A}_{\tilde{y}}$ and $\tilde{y} \in \mathcal{A}_{\tilde{x}}$. The matching surplus is assumed to be split by Nash bargaining. Let $\pi(\tilde{x}|\tilde{y})$ and $\pi(\tilde{y}|\tilde{x})$ be the endogenous flow payoff man $\tilde{x}$ and woman $\tilde{y}$ get from their matching. Then, we assume $\pi(\tilde{x}|\tilde{y}) + \pi(\tilde{y}|\tilde{x}) = f(\tilde{x}, \tilde{y})$. That is, there is no cost in splitting the matching surplus $f(\tilde{x}, \tilde{y})$.

When man $\tilde{x}$ and woman $\tilde{y}$ are matched, they do nothing except enjoy their flow payoff $\pi(\tilde{x}|\tilde{y})$ and $\pi(\tilde{x}|\tilde{y})$. We assume the matching surplus $f(\tilde{x}, \tilde{y})$ remains constant in a matching, so that matched agents will not separate voluntarily in a stationary equilibrium. To stabilize the unmatched pool, we assume matched agents has an exogenous separating rate δ. That is, if agents are matched at time t, the probability that they are still matched at time $t + \Delta t$ is $\exp(-\delta \cdot \Delta t)$. Finally, agents cannot search for other matchings on the market when they have already matched.

Let $U(\tilde{x})$ and $V(\tilde{y})$ denote the expected utility of unmatched man $\tilde{x}$ and woman $\tilde{y}$ with respect to the random meeting opportunities. Let $W(\tilde{x}|\tilde{y})$ and $W(\tilde{y}|\tilde{x})$ denote the expected utility of man $\tilde{x}$ and woman $\tilde{y}$ with respect to the random separation, respectively, when they are matched.

Since agents receive flow payoff π when matched, and are also subject to separation at Poisson rate δ. Thus, the Bellman equations for $W(\tilde{x}|\tilde{y})$ and $W(\tilde{y}|\tilde{x})$ are

$$
\begin{aligned}
rW(\tilde{x}|\tilde{y}) &= \pi(\tilde{x}|\tilde{y}) + \delta\left(U(\tilde{x}) - W(\tilde{x}|\tilde{y})\right) \\
rW(\tilde{y}|\tilde{x}) &= \pi(\tilde{y}|\tilde{x}) + \delta\left(V(\tilde{y}) - W(\tilde{y}|\tilde{x})\right)
\end{aligned} \tag{1}
$$

We assume the matching surplus is shared according to Nash bargaining. That is, $\pi(\tilde{x}|\tilde{y})$ and $\pi(\tilde{y}|\tilde{x})$ solve the following maximization problem,

$$
\begin{aligned}
\max_{\pi(\tilde{x}|\tilde{y}), \pi(\tilde{y}|\tilde{x})} & \left(W(\tilde{x}|\tilde{y}) - U(\tilde{x})\right)^{\theta} \left(W(\tilde{y}|\tilde{x}) - V(\tilde{x})\right)^{1-\theta} \\
\text{s.t. } & \pi(\tilde{x}|\tilde{y}) + \pi(\tilde{y}|\tilde{x}) = f(\tilde{x}, \tilde{y})
\end{aligned}
$$

where $\theta \in (0, 1)$ is the bargaining power of men. Combine with equation (1), we obtain

$$
\begin{aligned}
W(\tilde{x}|\tilde{y}) - U(\tilde{x}) &= \theta \cdot \frac{f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y})}{r + \delta} \\
W(\tilde{y}|\tilde{x}) - V(\tilde{y}) &= (1 - \theta) \cdot \frac{f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y})}{r + \delta}
\end{aligned} \tag{2}
$$

When unmatched, agents receive flow payoff $b(\tilde{x})$ or $c(\tilde{y})$, have chances to meet other unmatched agents, and form a match when both sides agree. Let $\tilde{\mathcal{G}}_t$ and $\tilde{\mathcal{Q}}_t$ denote the endogenous measure on characteristics of unmatched men and women at time t, respectively.

That is, for any set $A \subset \mathbb{R}^k$ and $B \subset \mathbb{R}^k$,

$$\tilde{\mathcal{G}}_t(A) = \mathbb{G}(\{i \in I : \tilde{x}_i \in A, \text{ and man } i \text{ is unmatched at time } t\}),$$
$$\tilde{\mathcal{Q}}_t(B) = \mathbb{Q}(\{j \in J : \tilde{y}_j \in B, \text{woman } j \text{ is unmatched at time } t\})$$

In stationary equilibrium, $\tilde{\mathcal{G}}_t$ and $\tilde{\mathcal{Q}}_t$ does not depends on time t. So, we can omit the time index t, and denote these distributions as $\tilde{\mathcal{G}}$ and $\tilde{\mathcal{Q}}$.

Thus, the Bellman equations for $U(\tilde{x})$ and $V(\tilde{y})$ are

$$\begin{aligned} rU(\tilde{x}) &= b(\tilde{x}) + \alpha \int_{\mathcal{A}_{\tilde{x}}} \mathbb{1}\left(\tilde{y} \in \mathcal{A}_{\tilde{y}}\right) \left(W(\tilde{x}|\tilde{y}) - U(\tilde{x})\right) \mathrm{d}\tilde{\mathcal{Q}}(\tilde{y}) \\ rV(\tilde{y}) &= c(\tilde{y}) + \alpha \int_{\mathcal{A}_{\tilde{y}}} \mathbb{1}\left(\tilde{y} \in \mathcal{A}_{\tilde{x}}\right) \left(W(\tilde{y}|\tilde{x}) - V(\tilde{y})\right) \mathrm{d}\tilde{\mathcal{G}}(\tilde{x}) \end{aligned} \tag{3}$$

where $\alpha \equiv m(u, v)/(uv)$.

A weakly dominant strategy for each agent is to set $\mathcal{A}_{\tilde{x}} = \{\tilde{y} : W(\tilde{x}|\tilde{y}) - U(\tilde{x}) \geq 0\}$ and $\mathcal{A}_{\tilde{y}} = \{\tilde{x} : W(\tilde{y}|\tilde{x}) - V(\tilde{y}) \geq 0\}$. Under this optimal behavior, equation (3) can be rewritten as

$$\begin{aligned} rU(\tilde{x}) &= b + \frac{\alpha\theta}{r+\delta} \int \max\{f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y}), 0\} \mathrm{d}\tilde{\mathcal{Q}}(\tilde{y}) \\ rV(\tilde{y}) &= c + \frac{\alpha(1-\theta)}{r+\delta} \int \max\{f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y}), 0\} \mathrm{d}\tilde{\mathcal{G}}(\tilde{x}) \end{aligned} \tag{4}$$

In a stationary economy, the flow from the matched to the unmatched should equal to the flow in the opposite direction. Thus, for any measurable set A and B in $\mathbb{R}^k$, we have

$$\begin{aligned} \delta\left(\mathcal{G}(A) - \tilde{\mathcal{G}}(A)\right) &= \alpha \int_A \int \mathbb{1}(f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y}) \geq 0) \mathrm{d}\tilde{\mathcal{Q}}(\tilde{y}) \mathrm{d}\tilde{\mathcal{G}}(\tilde{x}) \\ \delta\left(\mathcal{Q}(B) - \tilde{\mathcal{Q}}(B)\right) &= \alpha \int_B \int \mathbb{1}(f(\tilde{x}, \tilde{y}) - rU(\tilde{x}) - rV(\tilde{y}) \geq 0) \mathrm{d}\tilde{\mathcal{G}}(\tilde{x}) \mathrm{d}\tilde{\mathcal{Q}}(\tilde{y}) \end{aligned} \tag{5}$$

where the left-handed side denotes the measure of agents who just (exogenously) separated, and the right-handed side denotes the measure of agents who just got matched.

Definition 1. A *stationary equilibrium* (SE) can be represented as a tuple $(U, V, \tilde{\mathcal{G}}, \tilde{\mathcal{Q}}, u, v)$, such that (i) all agents behave optimally in accepting others so that equation (4) holds; (ii) flows into and out of the unmatched are balanced so that equation (5) holds, $u = \tilde{\mathcal{G}}(\mathbb{R}^k)$ and $v = \tilde{\mathcal{Q}}(\mathbb{R}^k)$.

4 Econometric Identification Problem

4.1 Observed and unobserved characteristics

We assume only part of the characteristics of agents is observable to economists. For example, the education level of agents is observable in most applications, yet the heterogeneous "ability" of agents with the same education level are unobserved in general. For the purpose of identification, we need to distinguish the observed and unobserved characteristics.

Assume the characteristics $\tilde{x}_i$ of man i and $\tilde{y}_j$ of woman j can be partitioned as $\tilde{x}_i = (x_i, \eta_i)$ and $\tilde{y}_j = (y_j, \zeta_j)$, where characteristics x_i and y_j are observable to economists but economists cannot observe characteristics η_i and ζ_j. Although x_i and y_j are observables and η_i and ζ_j are unobservables, all of them are agent's characteristics, and are common knowledge among agents.

For simplicity, we assume all observable characteristics $\{x_i, y_j\}$ are discrete. We call woman j belongs to group y if and only if $y_j = y$. Similarly, man i belongs to group x if and only if $x_i = x$. Let X and Y denote the set of all possible groups (observed characteristics) of men and women respectively. Throughout the paper, we assume the set X and Y are finite and use $|X|$ and $|Y|$ to denote the number of groups in X and Y respectively. In each group, there are infinitely many agents. Let μ_x and μ_y denote the total measure of agents in group x and y. By definition,

$$\mu_x = \mathcal{G}(\{(x', \eta') \in X \times \mathbb{R} : x' = x\})$$
$$\mu_y = \mathcal{Q}(\{(y', \zeta') \in Y \times \mathbb{R} : y' = y\})$$

We assume unobserved characteristics η and ζ are continuous variables and belongs to $\mathbb{R}$. Let G_x and Q_y denote the cumulative distribution function of η in group x and ζ in group y, respectively. By definition, for any η and ζ in $\mathbb{R}$,

$$G_x(\eta) = \frac{1}{\mu_x}\mathcal{G}(\{(x', \eta') \in X \times \mathbb{R} : x' = x, \eta' \leq \eta\})$$
$$Q_x(\zeta) = \frac{1}{\mu_y}\mathcal{Q}(\{(y', \zeta) \in Y \times \mathbb{R} : y' = y, \zeta' \leq \zeta\})$$

4.2 Separable matching surplus assumption

Our key identification assumption is to assume the matching surplus $f(\tilde{x}, \tilde{y})$ is separable in observable and unobserved characteristics. Moreover, we assume the matching surplus is also subject to a matching specific random shock.

Assumption 1 (Separable Matching Surplus)**.** *For any man with characteristics $\tilde{x} = (x, \eta)$ and any woman with characteristics $\tilde{y} = (y, \zeta)$, their matching surplus equals to*

$$f(\tilde{x}, \tilde{y}) = f_{xy} + \eta + \zeta + \varepsilon$$

where $f_{xy} \in \mathbb{R}$ is a (group-pairwise) matching surplus, and ε is a matching specific random shock independently drawn from distribution F_{xy} after man and woman meet. Moreover, ε is common knowledge and realized before agents make matching decisions.

The (group-pairwise) matching surplus matrix $\{f_{xy} : x \in X, y \in Y\}$ is the key parameter we would like to identify in the remainder of the paper. Assumption 1 is a natural extension of the separability assumption used in Choo and Siow (2006) and Galichon and Salanié (2014) in the context of matching model with time-consuming search frictions, except we assume the unobserved characteristics η and ζ are scalars instead of vectors.

What is actually assumed in Assumption 1 is the separability structure on the matching surplus and the additive structure of matching specific shock. It imposes no more restrictions than assuming

$$f(\tilde{x}, \tilde{y}) = f_{xy} + f_x(\eta) + f_y(\zeta) + \varepsilon,$$

where $\tilde{x} = (x, \eta)$ and $\tilde{y} = (y, \zeta)$. As we can then define $\hat{\eta} = f_x(\eta)$ and $\hat{\zeta} = f_y(\zeta)$, so that the matching surplus can be rewritten as $f_{xy} + \hat{\eta} + \hat{\zeta} + \varepsilon$, with $\hat{\eta}$ and $\hat{\zeta}$ distributed as $G_x \circ \hat{f}_x^{-1}$ and $Q_y \circ \hat{f}_y^{-1}$ respectively.

What such separability structure implies is that all complementarity (or, substitutability) of matching surplus is among observable characteristics x and y, so that unobserved types η and ζ shift agent's individual productivity or attractiveness but do not interact with the characteristics of his/her partners. As discussed in section 4.3 and 6, this leads to the nonparametric identification of (group-pairwise) matching surplus f_{xy} under very mild data requirement.

Finally, although both unobserved characteristics η, ζ and matching specific random shock ε perturb the matching surplus $f(\tilde{x}, \tilde{y})$ around its group benchmark f_{xy}, they have very different implications. ζ_j (η_i) is one part of woman j's (man i's) characteristics, therefore it will remain the same in all matchings participated by woman j (man i). To the contrary, ε_{ij} is drawn independently after man i and woman j meet and is specific to this match. In particular, the matching specific shock ε is independent of agent's characteristics η and ζ, though the distribution F_{xy} of ε can depend on observed characteristics x and y.

Throughout the paper, we keep the following regularity conditions.

Assumption 2 (Regularity Conditions)**.** *For each group $x \in X$ and group $y \in Y$, F_{xy}, G_x and Q_y have finite means. Moreover,*

(i) At least one of the following is true: (a) the support of F_{xy} has nonempty interior; (b) the support of G_x and Q_y have nonempty interior.

(ii) the support of F_{xy}, G_x and Q_y are (possibly infinite) intervals.

(iii) F_{xy}, G_x and Q_y admit positive, bounded and continuous density functions whenever their supports have nonempty interior.

We now rephrase the model in Section 3 in terms of observable and unobserved characteristics. Let $U_x(\eta)$ and $V_y(\zeta)$ denote the expected payoff of man (x, η) and woman (y, ζ) when unmatched. Let u_x and v_y denote the measure of unmatched agents in group x and y respectively. By definition, $u = \sum_x u_x$ and $v = \sum_y v_y$, and

$$
\begin{aligned}
u_x &= \tilde{\mathcal{G}}(\{(x', \eta') \in X \times \mathbb{R} : x' = x\}) \\
u_y &= \tilde{\mathcal{Q}}(\{(y', \zeta') \in Y \times \mathbb{R} : y' = y\})
\end{aligned}
$$

Assume agents have group specific flow payoff $\{b_x, c_y\}$ when unmatched. Then, the equilibrium condition (4) can be rewritten as

$$
\begin{aligned}
rU_x(\eta) &= b_x + \frac{\alpha\theta}{r+\delta} \sum_y v_y \iint \max(\psi_{xy}(\varepsilon, \eta, \zeta), 0) \, \mathrm{d}F_{xy}(\varepsilon) \, \mathrm{d}\tilde{Q}_y(\zeta) \\
rV_y(\zeta) &= c_y + \frac{\alpha(1-\theta)}{r+\delta} \sum_x u_x \iint \max(\psi_{xy}(\varepsilon, \eta, \zeta), 0) \, \mathrm{d}F_{xy}(\varepsilon) \, \mathrm{d}\tilde{G}_x(\eta)
\end{aligned} \tag{6}
$$

where

$$
\psi_{xy}(\varepsilon, \eta, \zeta) \equiv f_{xy} + \eta + \zeta + \varepsilon - rU_x(\eta) - rV_y(\zeta) \tag{7}
$$

is the flow of matching surplus net of reservation values, and $\tilde{G}_x(\cdot)$ and $\tilde{Q}_y(\cdot)$ are the endogenous distribution among the unmatched agents in group x and y respectively. In other words, they are the distribution of η and ζ conditional on unmatched agents. By definition,

$$
\begin{aligned}
\tilde{G}_x(\eta) &= \frac{1}{u_x} \tilde{\mathcal{G}}(\{(x', \eta') \in X \times \mathbb{R} : x' = x, \eta' \leq \eta\}) \\
\tilde{Q}_x(\zeta) &= \frac{1}{v_y} \tilde{\mathcal{Q}}(\{(y', \zeta) \in Y \times \mathbb{R} : y' = y, \zeta' \leq \zeta\}).
\end{aligned}
$$

Let $g_x(\cdot)$ and $\tilde{g}_x(\cdot)$ be the density function of $G_x(\cdot)$ and $\tilde{G}_x(\cdot)$, and $q_y(\cdot)$ and $\tilde{q}_y(\cdot)$ are the density functions of $Q_y(\cdot)$ and $\tilde{Q}_y(\cdot)$ respectively. Then, the flow condition (5) can be written as

$$\begin{aligned}\delta\left(g_x(\eta)\mu_x - \tilde{g}_x(\eta)u_x\right) = \tilde{g}_x(\eta)u_x \cdot \alpha\sum_y v_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right)\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}\tilde{Q}_y(\zeta)\\ \delta\left(q_y(\zeta)\mu_y - \tilde{q}_y(\zeta)v_y\right) = \tilde{q}_y(\zeta)v_y \cdot \alpha\sum_x u_x \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right)\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}\tilde{G}_x(\eta)\end{aligned} \tag{8}$$

Lemma 1 ensures that the endogenous distribution $\tilde{G}_x$ and $\tilde{Q}_y$ is well defined by equation (8). The proof is left in Appendix A.

Lemma 1. *Suppose Assumption 2 holds. Given function* $\{\psi_{xy}(\varepsilon,\eta,\zeta)\}$, $\{u_x, v_y\}$ *and* $\{\mu_x,\mu_y\}$, *there exists a unique collection of distributions* $\{\tilde{G}_x,\tilde{Q}_y : x\in X, y\in Y\}$ *which satisfy equation* (8). *Moreover, for each* $x\in X$ *and* $y\in Y$, *we have* $\mathrm{supp}(\tilde{G}_x) = \mathrm{supp}(G_x)$, $\mathrm{supp}(\tilde{Q}_y) = \mathrm{supp}(Q_y)$, *and both* $\tilde{G}_x$ *and* $\tilde{Q}_y$ *admit bounded and continuous density functions.*

After integrating equation (8), we get the following stationary condition for the measure of unmatched agents, $\{u_x, v_y\}$,

$$\begin{aligned}\delta(\mu_x - u_x) =& u_x\alpha\sum_y v_y \iiint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right)\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}\tilde{G}_x(\eta)\mathrm{d}\tilde{Q}_y(\zeta)\\ \delta(\mu_y - u_y) =& v_y\alpha\sum_x u_x \iiint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right)\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}\tilde{G}_x(\eta)\mathrm{d}\tilde{Q}_y(\zeta).\end{aligned} \tag{9}$$

After introducing observed and unobserved characteristics, we can rephrase the definition of stationary equilibrium in Definition 1 in terms of observed and unobserved characteristics.

Definition 2. A *stationary equilibrium* (SE) can be represented as a tuple $\{U_x, V_y, \tilde{G}_x, \tilde{Q}_y, u_x, v_y : x\in X, y\in Y\}$ such that (i) all agents behave optimally in accepting others so that equation (6) holds; (ii) flows into and out of the unmatched in each group are balanced so that equation (8) and (9) holds.

Shimer and Smith (2000) have proven the existence of SE in a setting with quadratic matching technology $m(\cdot,\cdot)$, and one can construct a proof based on Tröger and Nöldeke (2009) when $m(\cdot,\cdot)$ displays constant return to scale. However, these proofs assume (i) mathcing surplus function must be strict super-modular or sub-modular, (ii) no match-specific shocks ε on matching surplus, (iii) smooth distribution over agent's characteristics so that no discrete group is allowed. Therefore, the existing proof cannot be applied directly to our framework. For completeness, we provide an existence proof of SE in Appendix B adapted to the model discussed here.

Finally, Assumption 1 can also be interpreted as an assumption on agent's flow payoff when unmatched. To see this, assume flow payoffs when agents being unmatched, $\{b_x(\cdot), c_y(\cdot)\}$, are now functions of their unobserved characteristics, yet their matching surplus matrix $\{f_{xy}\}$ only depends on their observed characteristics. Then, the equilibrium condition (6) can be rewritten as

$$rU_x(\eta) = b_x(\eta) + \frac{\alpha\theta}{r+\delta} \sum_y v_y \iint \max\left(f_{xy} + \varepsilon - rU_x(\eta) - rV_y(\zeta), 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{Q}_y(\zeta)$$
$$rV_y(\zeta) = c_y(\zeta) + \frac{\alpha(1-\theta)}{r+\delta} \sum_x u_x \iint \max\left(f_{xy} + \varepsilon - rU_x(\eta) - rV_y(\zeta), 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{G}_x(\eta)$$

which is equivalent to

$$r\hat{U}_x = \frac{\alpha\theta}{r+\delta} \sum_y v_y \iint \max\left(\hat{\psi}(\varepsilon, \eta, \zeta), 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{Q}_y(\zeta)$$
$$r\hat{V}_y(\zeta) = \frac{\alpha(1-\theta)}{r+\delta} \sum_x u_x \iint \max\left(\hat{\psi}(\varepsilon, \eta, \zeta), 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{G}_x(\eta)$$

where $r\hat{U}_x(\eta) = rU_x(\eta) - b_x(\eta)$, $r\hat{V}_y(\zeta) = rV_y(\zeta) - c_y(\zeta)$ and

$$\hat{\psi}(\varepsilon, \eta, \zeta) = f_{xy} - b_x(\eta) - c_y(\zeta) - r\hat{U}_x(\eta) - r\hat{V}_y(\zeta).$$

In this way, it has exactly the same separable structure as in Assumption 1.

4.3 Data requirement

In this subsection, we discuss what data is needed to identify the group-pairwise matching surplus matrix $\{f_{xy}\}$ in Assumption 1.

First of all, we only need data from one single market. That is, our identification does not rely on variations between markets. Instead, we assume there is one single large market, and our approach is to find an invertible linkage between the data generated from a stationary equilibrium and the model primitives. Here, group-pairwise matching surplus $\{f_{xy}\}$ is the model primitives of interest. In this sense, this paper is different from Fox (2010) and Fox et al. (2015) , and share the same idea with Choo and Siow (2006), Galichon and Salanié

(2014) and Fox and Bajari (2013) etc.

Group-pairwise matching patterns Let μ_{xy} denote the total measure of the matched agents between group x and group y. We call $\{\mu_{xy}\}$ as *group-pairwise matching patterns* throughout the paper, and abbreviate it as *matching patterns* if there is no confusion in the context. By definition, we have accounting equalities $\mu_x = u_x + \sum_y \mu_{xy}$ and $\mu_y = v_y + \sum_x \mu_{xy}$ for all group x and y. Therefore, u_x and v_y can be calculated from matching patterns $\{\mu_{xy}\}$, and measures of agents in each group $\{\mu_x, \mu_y\}$. We treat $\{\mu_x, \mu_y\}$ as known.

In a stationary equilibrium, the measure of separated matching between group x and y should equal to the measure of newly matched pairs from group x and y. That is, the following flow condition must hold,

$$\delta\mu_{xy} = \alpha u_x v_y \iiint \mathbb{1}\left(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{G}_x\left(\eta\right) \mathrm{d}\tilde{Q}_y\left(\zeta\right) \tag{10}$$

where the left-handed side is the flow from the matched pairs between group x and y to the unmatched pool, and the right-handed side is the measure of newly formed matching from unmatched agents in group x and group y. Note $\alpha u_x v_y$ is the flow rate in which agents in group x and y meet each other in the random search process, while the integral of the indicator function describes the ratio of successful matchings per meeting.

As can be seen from equation (10), the group-pairwise matching surplus $\{f_{xy}\}$ affects the matching pattern $\{\mu_{xy}\}$ in three ways: (*i*) It affects the acceptance probability by directly changing the matching surplus through the group mean effect; (*ii*) It influence the acceptance probability indirectly by changing the reserve value of each agents, as indicated in equation (6); (*iii*) It influence the distribution of agents an individual will meet in the unmatched pool, as indicated in the flow condition (8).

As in the frictionless matching literature[1], matching pattern $\{\mu_{xy}\}$ often contains information on the group-pairwise matching surplus $\{f_{xy}\}$. Unlike the frictionless matching literature, however, the presence of endogenous conditional distribution and the interaction between equation (6), (8) and (10) make it more difficult to solve for $\{f_{xy}\}$ directly from $\{\mu_{xy}\}$. To achieve identification in a more transparent way, we utilize the distribution of unmatched duration across agents, in addition to the matching pattern.

Distribution of unmatched duration For any unmatched man (x, η) , the probability that the length of his unmatched duration is greater than l is $\exp(-l \cdot h_x(\eta))$ where $h_x(\eta)$ is

[1] See, for example, Galichon and Salanié (2014) and Choo and Siow (2006).

the hazard rate of his unmatched duration,

$$h_x(\eta) = \alpha \sum_y v_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta) \geq 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{Q}_y(\zeta). \tag{11}$$

Note the flow rate for a man (x, η) meeting someone in group y is αv_y, and the chance of getting matched in a meeting is

$$\iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta) \geq 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{Q}_y(\zeta).$$

Similarly, the hazard rate of woman (y, ζ)'s unmatched duration is

$$h_y(\zeta) = \alpha \sum_x u_x \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta) \geq 0\right) \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}\tilde{G}_x(\eta). \tag{12}$$

Given Lemma 1, both $h_x(\cdot)$ and $h_y(\cdot)$ are continuous under Assumption 2.

Let D_x and D_y be the cumulative distribution function of unmatched duration among unmatched agents in group x and group y respectively. Fix an arbitrary time t_0. For any duration length $l > 0$, $D_x(l)$, by definition, denotes the fraction of agents in group x who are unmatched at time t_0 and have got matched before time $t_0 + l$ since time t_0. That is,

$$\begin{aligned} D_x(l) &= \int (1 - e^{-l \cdot h_x(\eta)}) \mathrm{d}\tilde{G}_x(\eta) \\ D_y(l) &= \int (1 - e^{-l \cdot h_y(\zeta)}) \mathrm{d}\tilde{Q}_y(\zeta) \end{aligned} \tag{13}$$

Note the data needed to compute $\{D_x, D_y\}$ is less demanding than the long panel data requirement in the literature. In their simulation examples, Hagedorn et al. (2014) assumes a panel with 10-20 years at monthly frequency, and Lamadon et al. (2013) assumes a panel with 40 years at quarterly frequency. Yet, for our purpose, a 5-year[2] panel may be more than enough to estimate $\{D_x, D_y\}$ to a satisfactory precision in the labor market context. Moreover, $\{D_x, D_y\}$ may be constructed from a cross-sectional data in some cases. For example, most unemployment surveys ask unemployed workers to report how long they have been unemployed up to now. The distributions of such up-to-now unemployment duration are equivalent to distributions $\{D_x, D_y\}$ in stationary equilibrium. Although stationary equilibrium is a maintained assumption in our paper and all these papers, it may be a good

[2] For example, in the Current Population Survey(CPS) conducted by the Bureau of Labor Statistics (BLS) of United States, respondents were able to report unemployment durations of up to 2 years before 2011, and 5 years after 2011.

model approximation of the reality within a 5-year period, but a poor one over decades, as there might be long-term trend on model primitives. Finally, our mild data requirement is also helpful in dealing data attrition problem in empirical applications.

In the remainder of the paper, we always assume the matching pattern $\{\mu_{xy}, u_x, v_y\}$ and distribution of unmatched duration $\{D_x, D_y\}$ can be derived from the data. As shown in section 6, such data is sufficient to identify matching surplus $\{f_{xy}\}$ if distribution $\{G_x, Q_y, F_{xy}\}$ and other parameters are known.

4.4 Interior solution assumption

We should be careful that our key identification equation (10) on $\{\mu_{xy}\}$ could be degenerate. Consider the case of a corner solution when, for some $x^* \in X$ and $y^* \in Y$, $\psi_{x^*y^*}(\varepsilon, \eta, \zeta) \geq 0$ for any $(\varepsilon, \eta, \zeta)$ in the support. In this case, equation (10) on group x^* and y^* is reduced to

$$\delta\mu_{x^*y^*} = \alpha u_{x^*} v_{y^*} \tag{14}$$

and an inequality

$$\psi_{x^*y^*}(\varepsilon, \eta, \zeta) \geq 0, \forall (\varepsilon, \eta, \zeta) \in \operatorname{supp}(F_{xy}) \times \operatorname{supp}(G_x) \times \operatorname{supp}(Q_y). \tag{15}$$

Equation (14) implies the matching between men in group x and women in group y is only restricted to the search friction: they always get matched when they meet. In particular, such $\mu_{x^*y^*}$ contains no more information on $f_{x^*y^*}$ other than inequality (15). With inequality (15), we cannot pin down $rU_{x^*}(\cdot)$ in equation (6). In fact, $rU_{x^*}(\eta)$ could be any large number, which further undermines the identification of any other f_{x^*y} with $y \neq y^*$.

To restore identification, we make the following high-level assumption on $\{f_{xy}\}$.

Assumption 3 (Interior Solution)**.** *For any $x \in X$ and $y \in Y$, assume $\delta\mu_{xy} < \alpha u_x v_y$.*

Assumption 3 means agents in group x and y are "picky" so that not every meeting between them will result in a matching. In empirical work, Assumption 3 can be easily checked from the data. Moreover, for any data, there always exists a small enough δ such that Assumption 3 holds. The following lemma gives an sufficient condition for Assumption 3 based on the model primitives. The proof is in Appendix C.

Lemma 2. *Suppose Assumption 1 and 2 hold. Assumption 3 holds, if for each group $x \in X$ and group $y \in Y$, $f_{xy} + \underline{\varepsilon}_{xy} + \underline{\eta}_x + \underline{\zeta}_y < b_x + c_y$, where $\underline{\varepsilon}_{xy} = \inf \operatorname{supp}(F_{xy})$, $\underline{\eta}_x = \inf \operatorname{supp}(G_x)$ and $\underline{\zeta}_y = \inf \operatorname{supp}(Q_y)$.*

Since the flow payoff when agents being unmatched, b_x and c_y, sets the lower bound of rU_x and rV_y, lemma 2 means Assumption 3 holds if the least capable agents in group x and y will never get matched when they have the least favorable matching specific random shock. Such restriction on the parameter space of $\{f_{xy}\}$ is sufficient for Assumption 3. In particular, Assumption 3 is trivially satisfied if one of G_x, Q_y and F_{xy} has no lower bound in its support.

There is another kind of corner solution $\mu_{xy} = 0$ which would leads to partial identification. We postpone its discussion to section 6.2.

5 Identification without Unobserved Characteristics

Suppose, for now, there is no unobserved characteristics. That is, $\text{supp}(G_x) = \{0\}$ and $\text{supp}(Q_y) = \{0\}$ for all x and y. And, we assume $\text{supp}(F_{xy}) \supset \mathbb{R}_+$ for simplicity. In this case, equation (6) can be rewritten as

$$\begin{aligned} rU_x &= b_x + \frac{\alpha\theta}{r+\delta} \sum_y v_y \int \max\left(f_{xy} + \varepsilon - rU_x - rV_y, 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \\ rV_y &= c_y + \frac{\alpha\left(1-\theta\right)}{r+\delta} \sum_x u_x \int \max\left(f_{xy} + \varepsilon - rU_x - rV_y, 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \end{aligned} \tag{16}$$

And, the flow condition (8) can now be simplified as

$$\begin{aligned} \delta\left(\mu_x - u_x\right) &= u_x \cdot \alpha \sum_y v_y \left(1 - F_{xy}\left(-f_{xy} + rU_x + rV_y\right)\right) \\ \delta\left(\mu_y - v_y\right) &= v_y \cdot \alpha \sum_x u_x \left(1 - F_{xy}\left(-f_{xy} + rU_x + rV_y\right)\right) \end{aligned} \tag{17}$$

Finally, the matching patterns in equation (10) is

$$\delta\mu_{xy} = \alpha u_x v_y \left(1 - F_{xy}\left(-f_{xy} + rU_x + rV_y\right)\right) \tag{18}$$

Compared to the original conditions, the assumption here simplifies the problem in two ways: (i) Since all agents are now assumed to be homogeneous within each group, the reservation value U_x and V_y is now just a real number instead of an infinite-dimensional object; (ii) Since there is no unobserved characteristics, we don't need to worry about the endogenous conditional distribution $\{\tilde{G}_x, \tilde{Q}_y\}$, which are degenerate as the unconditional distribution $\{G_x, Q_y\}$.

Because of these two simplification, as long as $0 < \delta\mu_{xy} < \alpha u_x v_y$, we can invert equation (18) as

$$f_{xy} - rU_x - rV_y = -F_{xy}^{-1}\left(1 - \frac{\delta\mu_{xy}}{\alpha u_x v_y}\right). \tag{19}$$

Then, U_x and V_y can be solved from (16) and (19). Therefore, group-pairwise matching surplus f_{xy} can be identified by the matching pattern $\{\mu_{xy}\}$, given other parameters.

As discussed in Chiappori and Salanié (2015), frictionless matching models can rationalize all possible matching patterns. That is, for any matching pattern $\{\mu_{xy}\}$ satisfying

$$\sum_y \mu_{xy} < \mu_x,\ \sum_x \mu_{xy} < \mu_y,\ \mu_x > 0,\ \mu_y > 0,\ \forall x \in X, y \in Y, \tag{20}$$

there exists a matching surplus matrix $\{f_{xy}\}$ such that $\{\mu_{xy}\}$ is consistent with its frictionless matching equilibrium.

In the frictional model discussed here, one needs an extra restriction for a rationalizable matching pattern: $\delta\mu_{xy} \le \alpha u_x v_y$, for all x and y. That is, the matching pattern should be rationalizable as a flow equilibrium. For a very large separation rate δ, the model cannot support a large μ_{xy} simply because the speed of generating a matching is upper bounded by the time-consuming search friction. Even if every meeting successfully results in a matching, agents need time to meet each other.

On the other hand, however, for any matching pattern $\{\mu_{xy}\}$ satisfying condition (20), there exists some small enough separation rate $\delta > 0$ and large enough meeting rate $\alpha > 0$, such that this matching pattern $\{\mu_{xy}\}$ can be rationalized by the model.

We summarize the above discussion in the following proposition.

Proposition 1. *Suppose Assumption 1, 2 and 3 hold,* $\text{supp}\,(G_x) = \text{supp}\,(Q_y) = \{0\}$ *and* $\text{supp}(F_{xy}) \supset \mathbb{R}_+$ *for all* $x \in X$ *and* $y \in Y$*. Then,* $\{f_{xy} : x \in X, y \in Y\}$ *can be identified in the parameter space* $\mathbb{R}^{|X|\cdot|Y|}$ *by matching patterns* $\{\mu_{xy}\}$*, given* $\{F_{xy} : x \in X, y \in Y\}$*, interest rate* r*, bargaining power* θ*, separation rate* δ*, matching rate* α *and agent's flow payoff* $\{b_x, c_y\}$ *when unmatched.*

Moreover, for any $\{\mu_{xy}\}$ *satisfying condition (20), there exists some* $\delta^* > 0$ *and* $\alpha^* > 0$*, such that for any* $\delta \in (0, \delta^*)$ *and* $\alpha \in (\alpha^*, +\infty)$*, such* $\{\mu_{xy}\}$ *is rationalizable.*

6 Identification with Unobserved Characteristics

In this section, we deal with the identification problem with unobserved characteristics of agents. As discussed in section 4.3, we assume both matching pattern $\{\mu_{xy}\}$ and distribution of unmatched duration $\{D_x, D_y\}$ is observable.

We divide the identification into two steps. In the first step, we identify the conditional distribution $\tilde{G}_x$ and $\tilde{Q}_y$ from $\{D_x, D_y\}$. Next, we use equation (10) to identify the matching surplus $\{f_{xy}\}$ from the matching patterns $\{\mu_{xy}\}$.

6.1 Identifying conditional distributions $\{\tilde{G}_x, \tilde{Q}_y : x \in X, y \in Y\}$

let's start from the observation that the conditional distribution $\{\tilde{G}_x, \tilde{Q}_y\}$ of unmatched agents can be rewritten as a function of unconditional distribution $\{G_x, Q_y\}$ and agent's hazard rate of unmatched duration $\{h_x, h_y\}$ defined in (11) and (12). Substitute equations (11) and (12) into the flow condition (8), we have

$$\begin{aligned} \tilde{g}_x(\eta) &= \frac{\delta \mu_x}{(\delta + h_x(\eta))\, u_x} g_x(\eta) \text{ a.e.} \\ \tilde{q}_y(\zeta) &= \frac{\delta \mu_y}{(\delta + h_y(\zeta))\, v_y} q_y(\zeta) \text{ a.e.} \end{aligned} \tag{21}$$

Thus, as long as we can identify hazard rate $h_x(\cdot)$ and $h_y(\cdot)$, we can deduce the endogenous distribution $\tilde{G}_x$ and $\tilde{Q}_y$ from the exogenous distribution G_x and Q_y.

Now, it's time to explore the separable structure in Assumption 1. Equation (11) implies that, $h_x(\eta_1) \leq h_x(\eta_2)$ if $\eta_1 - rU_x(\eta_1) \leq \eta_2 - rU_x(\eta_2)$. Similarly, $h_y(\zeta_1) \leq h_y(\zeta_2)$ if $\zeta_1 - rV_y(\zeta_1) \leq \zeta_2 - rV_y(\zeta_2)$.

Taking derivative of $rU_x(\eta)$ and $rV_y(\zeta)$ in equation (6), we have for any $\eta \in \text{supp}(G_x)$ and $\zeta \in \text{supp}(Q_y)$.

$$1 - \frac{\partial rU_x(\eta)}{\partial \eta} = \frac{1}{1 + \frac{\theta}{r+\delta} h_x(\eta)} > 0 \tag{22}$$

$$\text{and similarly, } 1 - \frac{\partial rV_y(\zeta)}{\partial \zeta} = \frac{1}{1 + \frac{1-\theta}{r+\delta} h_y(\zeta)} > 0, \tag{23}$$

which implies $\eta - rU_x(\eta)$ and $\zeta - rV_y(\zeta)$ is a strictly increasing function of η and ζ. Therefore, the hazard rate $h_x(\eta)$ and $h_y(\zeta)$ are non-decreasing in η and ζ within each group.

Combine (13) and (21) , we have, for any $l > 0$, that

$$\int \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta)\mathrm{d}\eta = \frac{u_x(1 - D_x(l))}{\delta\mu_x}$$
$$\int \frac{e^{-lh_y(\zeta)}}{\delta + h_y(\zeta)} q_y(\zeta)\mathrm{d}\zeta = \frac{v_y(1 - D_y(l))}{\delta\mu_y}$$

Heuristically, suppose $\mathrm{supp}(F_{xy}) = \mathrm{supp}(G_x) = \mathrm{supp}(Q_y) = \mathbb{R}$ so that $h_x(\eta)$ and $h_y(\zeta)$ is strictly increasing and continuously differentiable. We can then rewrite the above equation by changing integrated variable from η and ζ to h,

$$\int e^{-lh}\lambda_x(h)\mathrm{d}h = \frac{u_x(1 - D_x(l))}{\delta\mu_x}$$
$$\int e^{-lh}\lambda_y(h)\mathrm{d}h = \frac{v_y(1 - D_y(l))}{\delta\mu_y}$$

where

$$\lambda_x(h) = \frac{g_x(h_x^{-1}(h))}{(\delta + h) \cdot h'_x(h_x^{-1}(h))}$$
$$\lambda_y(h) = \frac{q_y(h_y^{-1}(h))}{(\delta + h) \cdot h'_y(h_y^{-1}(h))}$$

Since $\int \exp(-l \cdot h)\lambda_x(h)\mathrm{d}h$ can be viewed as the Laplace transformation[3] of $\lambda_x(\cdot)$, and since Laplace transformation is invertible, we can identify $\lambda_x(h)$ and, thus, $g_x(h_x^{-1}(h))/h'_x(h_x^{-1}(h))$ from $D_x(\cdot)$. Note

$$G_x(h_x^{-1}(h)) = \int_0^h \frac{g_x(h_x^{-1}(h))}{h'_x(h_x^{-1}(h))}\mathrm{d}h$$

from which we can identify $h_x^{-1}(h)$ and, hence, $h_x(\eta)$. We can identify $h_y(\zeta)$ in a similar way.

Therefore, we have proved the following proposition when $h_x(\eta)$ and $h_y(\zeta)$ is differentiable and invertible. As shown in Appendix D, the result holds under more general cases.

Proposition 2. *Suppose Assumption 1 and 2 hold, and, for any $x \in X$ and $y \in Y$, the support of F_{xy}, G_x and Q_y have nonempty interior. Then, the endogenous conditional distribution $\{\tilde{G}_x, \tilde{Q}_y : x \in X, y \in Y\}$ and the hazard rate of unmatched duration $\{h_x(\cdot), h_y(\cdot) : x \in X, y \in Y\}$ can be identified from the distribution of unmatched duration $\{D_x, D_y : x \in X, y \in Y\}$, given $\{G_x, Q_y : x \in X, y \in Y\}$ and separation rate δ.*

[3] Standard Laplace transformation requires $\lambda_x(\cdot)$ to be a continuous function defined on $(0, +\infty)$, but the $\lambda_x(\cdot)$ in our case is a continuous function only defined on $(0, \alpha v)$. However, this nonstandard Laplace transformation is still invertible by lemma 7 in the appendix.

6.2 Identifying group-pairwise matching surplus $\{f_{xy} : x \in X, y \in Y\}$

Given the result in Proposition 2, it's now tractable to work with equations (10). We rewrite it as

$$\frac{\delta\mu_{xy}}{\alpha u_x v_y} = \iiint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta) \geq 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{G}_x\left(\eta\right) \mathrm{d}\tilde{Q}_y\left(\zeta\right) \tag{24}$$

The remaining challenge is now how to constructively identify $\{f_{xy}\}$ when the reservation value $\{rU_x(\cdot), rV_y(\cdot)\}$ are determined implicitly by $\{f_{xy}\}$ in condition (6).

Note that the slope of function $\psi_{xy}(\varepsilon,\eta,\zeta)$ is governed by the slope of $\eta - rU_x(\eta)$ and $\zeta - rV_y(\zeta)$ which we've already known from equation (22) and (23). Assume, without loss of generality, $(0,0,0)$ is in $\operatorname{supp}(F_{xy}) \times \operatorname{supp}(G_x) \times \operatorname{supp}(Q_y)$. Then, we can rewrite ψ_{xy} as

$$\psi_{xy}\left(\varepsilon,\eta,\zeta\right) = \psi_{xy}\left(0,0,0\right) + \varepsilon + \int_0^{\eta} \frac{1}{1+\frac{\theta}{r+\delta}h_x\left(\eta\right)} \mathrm{d}\eta' + \int_0^{\zeta} \frac{1}{1+\frac{1-\theta}{r+\delta}h_y\left(\zeta\right)} \mathrm{d}\zeta'. \tag{25}$$

This implies function $\psi_{xy}(\varepsilon,\eta,\zeta)$ is fully determined by a scalar $\psi_{xy}(0,0,0)$. As a result, we can view $\delta\mu_{xy}/(\alpha u_x v_y)$ in equation (24) as an increasing function of $\psi_{xy}(0,0,0)$.

Suppose $\delta\mu_{xy}/(\alpha u_x v_y) \in (0,1)$. Assumption 2 and Lemma 1 guarantee that $\delta\mu_{xy}/(\alpha u_x v_y)$ is continuous and strictly increasing in $\psi_{xy}(0,0,0)$. Since $\delta\mu_{xy}/(\alpha u_x v_y)$ can be calculated from the matching pattern $\{\mu_{xy}\}$ and model parameter $\{\mu_x, \mu_y\}$, we can identify $\psi_{xy}(0,0,0)$ and, hence, function ψ_{xy} from (24).

Suppose $\delta\mu_{xy} = 0$. We know $\psi_{xy}(\varepsilon,\eta,\zeta) \leq 0$ for any (ε,η,ζ) in the support, but we cannot point identify ψ_{xy}.

In both cases, however, we can calculate reservation value $\{rU_x(\cdot), rV_y(\cdot)\}$ from equation (6), since we now know either the exact value of function ψ_{xy}, or that $\psi_{xy}(\varepsilon,\eta,\zeta) \leq 0$ for any (ε,η,ζ) in the support, and since the endogenous distributions $\tilde{G}_x$ and $\tilde{Q}_y$ have been identified in Proposition 2.

Finally, the knowledge of $\{rU_x(\cdot), rV_y(\cdot)\}$ enable us to (partially) identify the group-pairwise matching surplus $\{f_{xy}\}$. For (x,y) such that $\mu_{xy} > 0$, f_{xy} can be point identified as

$$f_{xy} = \psi_{xy}\left(\varepsilon,\eta,\zeta\right) - \varepsilon - \left(\eta - rU_x\left(\eta\right)\right) - \left(\zeta - rV_y\left(\zeta\right)\right). \tag{26}$$

For (x,y) such that $\mu_{xy} = 0$, f_{xy} cannot be point identified, but we know from $\psi_{xy}(\varepsilon,\eta,\zeta) \leq$

0 that f_{xy} is in $(-\infty, \bar{f}_{xy}]$ where

$$\bar{f}_{xy} = (rU_x(\bar{\eta}) - \bar{\eta}) + (rV_y(\bar{\zeta}) - \bar{\zeta}) - \bar{\epsilon} \tag{27}$$

and $\bar{\varepsilon} = \sup \operatorname{supp}(F_{xy})$, $\bar{\eta} = \sup \operatorname{supp}(G_x)$, $\bar{\zeta} = \sup \operatorname{supp}(Q_y)$. Moreover, the upper bound $\bar{f}_{xy}$ is sharp in the sense that any $f_{xy} \leq \bar{f}_{xy}$ can rationalize the data we use as a stationary equilibrium.

Therefore, we have proven the following identification result constructively.

Proposition 3. *Suppose Assumption 1, 2, 3 and assumptions in Proposition 2 hold. Moreover, suppose distribution $\{G_x, Q_y, F_{xy}\}$, interest rate r, bargaining power θ, separation rate δ, matching rate α, agent's flow payoff $\{b_x, c_y\}$ and measures of groups $\{\mu_x, \mu_y\}$ are known.*

Then, group-pairwise matching surplus $\{f_{xy}\}$ is (partially) identified from matching patterns $\{\mu_{xy}\}$. In particular, for each $x \in X$ and $y \in Y$,

(i) f_{xy} is point identified as in equation (26) if $\mu_{xy} > 0$;

(ii) f_{xy} is set identified as in $(-\infty, \bar{f}_{xy}]$ if $\mu_{xy} = 0$. Moreover, $\bar{f}_{xy}$ defined in (27) is a sharp upper bound.

7 Discussion

In empirical research, matching and search problems in different context can have quite different structures. Although we choose to build our identification strategy upon the canonical matching-search-bargain framework as in Shimer and Smith (2000), it's likely that some modification is needed to carry our identification strategy into specific empirical questions. In this section, we discuss two extension/modification of our benchmark result.

7.1 Free entry

In the benchmark model, we assume the measure of men and women in each group $\{\mu_x, \mu_y\}$ is fixed. In some applications, however, agents enter into the market freely, and both $\{\mu_x, \mu_y\}$ and unconditional distribution $\{G_x, Q_y\}$ are determined endogenously by the free entry condition. How should this modification changes the economic model and econometric identification?

In terms of economic model, suppose initially there are $\bar{\mu}_x$ total measure of agents in group x and $\bar{\mu}_y$ in gruop y, and the initial distribution of unobserved characteristics is $\bar{G}_x$ in group x and $\bar{Q}_y$ in group y.

In equilibrium, there exists an endogenous cutoff η_x of unobserved characteristics in group x, and cutoff ζ_y in group y, such tbar for any x, y

$$rU_x(\eta_x) = b_x, \quad rV_y(\zeta_y) = c_y \tag{28}$$

and only agents with unobserved characteristics above cutoff η_x (or, cutoff ζ_y) are *active* and they will keep searching when they are unmatched. Those with characteristics lower than the cutoff will exit the market endogenously. Therefore, the measure of *active* agents in group x and y is endogenously determined as

$$\mu_x = \bar{\mu}_x(1 - \bar{G}_x(\eta_x)), \quad \mu_y = \bar{\mu}_y(1 - \bar{Q}_y(\zeta_y))$$

and the distributions of all *active* agents $\{G_x, Q_y\}$ are also endogenously determined as

$$G_x(\eta) = \mathbb{1}(\eta \geq \eta_x)\frac{\bar{G}_x(\eta) - \bar{G}_x(\eta_x)}{1 - \bar{G}_x(\eta_x)}, \quad Q_y(\zeta) = \mathbb{1}(\zeta \geq \zeta_y)\frac{\bar{Q}_y(\zeta) - \bar{Q}_y(\zeta_y)}{1 - \bar{Q}_y(\zeta_y)}$$

The free entry condition (28) along with previous equilibrium conditions (6), (8) and (9) defines the stationary equilibrium in this modified model.

In terms of econometric identification of matching surplus $\{f_{xy}\}$, however, little has been changed. The reasoning in section 6 still holds. From the econometric point of view, as long as the measure of active agents $\{\mu_x, \mu_y\}$ is still observable, whether it is exogenous or endogenous does not matter. The only problem is thar it's now unrealistic to assume distributions $\{G_x, Q_y\}$ are known to researchers, as these distributions are now themselves endogenous under free entry conditions. Therefore, we will discuss how distributions $\{G_x, Q_y\}$ can be identified from the data in the next subsection.

Finally, it's hard to construct counterfactuals in this model, because the initial distributions $\{\bar{G}_x, \bar{Q}_y\}$ are only partially identified. Even if we can identify distributions $\{G_x, Q_y\}$ as discussed in the next subsection, we have no information on $\bar{G}_x(\eta)$ and $\bar{Q}_y(\zeta)$ for $\eta < \eta_x$ and $\zeta < \zeta_y$, since the part below the cutoff is not revealed in the equilibrium. One solution is to assume *recoverability condition* as in Flinn and Heckman (1982), and another possible solution is to collect data from multiple markets.

7.2 Identification of distribution $\{G_x, Q_y\}$

One disadvantage of our identification result in section 6 is that it builds on the knowledge of distributions $\{G_x, Q_y\}$. There are two ways to identify distributions $\{G_x, Q_y\}$ in empirical applications.

One possible solution is to parameterize matching surplus $\{f_{xy}\}$ and distribution $\{G_x, Q_y\}$. Generally, the identification in a parameterized model depends on how $\{f_{xy}, G_x, Q_y\}$ is parameterized. As a result, the identification result should be studied on a case-by-case basis.

Instead of working out specific parametric examples, we focus on whether the likelihood function is well defined. Indeed, although Shimer and Smith (2000) and Proposition 5 in Appendix B ensure the existence of a stationary equilibrium, there could be multiple equilibria. This give rise to the question whether there exists a well-defined likelihood function mapping from the parameter space to the distributions of data.

Fortunately, the answer is affirmative when Proposition 2 holds. For a specific parameter, distributions $\{G_x, Q_y, F_{xy}\}$ are known, and Proposition 2 shows there is only one unique set of hazard rate functions $\{h_x, h_y\}$ compatible with the data on the distribution of unmatched duration $\{D_x, D_y\}$. Since both conditional distribution $\{\tilde{G}_x, \tilde{Q}_y\}$ and value function $\{rU_x, rV_y\}$ can be calculated from $\{h_x, h_y\}$, we know which equilibrium is actually realized in the data.

Another possible solution is to identify $\{G_x, Q_y\}$ nonparametrically by panel data on on micro-level matching patterns. Such panel data should have long enough time series observations on each agent.

To see how it works, suppose we have a panel data $\{m_{ijt}, x_i, y_j : i \in I, j \in J, t \in [0, T]\}$ on a continuum set of men I and women J from time 0 to T, where $m_{ijt} = 1$ if and only if man i is matched with woman j at time t, and x_i and y_j are the observed characteristics of man i and woman j, respectively. As in Section 3, i and j refers to the identify of each man and woman in the data, economists can observe their characteristics x_i and y_j but cannot observe characteristics η_i and ζ_j. Recall $\mathbb{G}$ and $\mathbb{Q}$ denote the measure on I and J. By definition, $u_x = \mathbb{G}(\{i : x_i = x, m_{ijt} = 0, \forall j\})$, $v_y = \mathbb{Q}(\{j : y_j = y, m_{ijt} = 0, \forall i\})$ at any time t. Note u_x and v_y remain the same at any time t in a stationary equilibrium.

Suppose T is large enough so that we can calculate the hazard rate h_i and h_j of unmatched duration for each agent i and j. Let ρ_i and ϱ_j be the quantile of man i and woman j in terms of h_i and h_j in group x_i and y_j. Formally,

$$\rho_i = \frac{\mathbb{G}(\{i' \in I : x_{i'} = x_i, h_{i'} \leq h_i\})}{\mathbb{G}(\{i' \in I : x_{i'} = x_i\})}$$
$$\varrho_j = \frac{\mathbb{Q}(\{j' \in J : y_{j'} = y_j, h_{j'} \leq h_j\})}{\mathbb{Q}(\{j' \in J : y_{j'} = y_j\})}.$$

Define $\tilde{G}_x^q(\rho)$ and $\tilde{Q}_y^q(\varrho)$ as the distribution of ρ and ϱ among unmatched agents. Formally, at any fixed time t,

$$\tilde{G}_x^q(\rho) = \frac{1}{u_x}\mathbb{G}(\{i : x_i = x, \rho_i \leq \rho, m_{ijt} = 0, \forall j\})$$
$$\tilde{Q}_y^q(\varrho) = \frac{1}{v_y}\mathbb{Q}(\{j : y_j = y, \varrho_j \leq \varrho, m_{ijt} = 0, \forall i\}).$$

Finally, for any $\rho, \varrho \in (0, 1)$, let $M_{xy}(\rho, \varrho)$ denote the measure of matched agents in group x and y with $\rho_i < \rho$ and $\varrho_j < \varrho$. Formally, at any time t, we have

$$M_{xy}(\rho, \varrho) = \iint m_{ijt}\mathbb{1}(x_i = x, y_j = y, \rho_i \leq \rho, \varrho_j \leq \varrho)\mathrm{d}\mathbb{G}(i)\mathrm{d}\mathbb{Q}(j)$$

Let μ_{xy}, $\tilde{g}_x^q$ and $\tilde{q}_y^q$ be the density of distribution M_{xy}, $\tilde{G}_x^q$ and $\tilde{Q}_y^q$. All these three densities can be identified from the data.

To identify $\{G_x, Q_y\}$, we need the following assumption.

Assumption 4. *Suppose, for any $x \in X$ and $y \in Y$,*

(i) $\operatorname{supp}(F_{xy}) \supset \mathbb{R}_-$*;*

(ii) for any $\eta > \underline{\eta}_x$, $rU_x(\eta) > rU_x(\underline{\eta}) = b_x$ where $\underline{\eta}_x = \inf \operatorname{supp}(G_x)$;

(iii) for any $\zeta > \underline{\zeta}_y$, $rV_y(\zeta) > rV_y(\underline{\zeta}) = c_y$ where $\underline{\zeta}_y = \inf \operatorname{supp}(Q_y)$.

The first condition in Assumption 4 implies for any man i and woman j, there exists some matching specific shock ε such that they are not willing to form a matching. As a result, $h_x(\eta) < \alpha v$ and $h_y(\zeta) < \alpha u$ for any η and ζ in the support. The second and third conditions in Assumption 4 are free entry conditions. As discussed in the preceding sections, we can never recover G_x and Q_y below the free entry cutoff, as we have little information on agents who are never engaged in any matching.

Assumption 4 implies that $h_x(\eta) \in (0, \alpha v)$ and $h_y(\zeta) \in (0, \alpha u)$ for any η and ζ in the interior of the support. Hence, $h_x(\eta)$ and $h_y(\zeta)$ are strictly increasing in η and ζ.

In stationary equilibrium, we have

$$\begin{aligned}\delta\mu_{xy}(\rho,\varrho) &= \alpha u_x v_y \cdot \tilde{g}^q_x(\rho)\cdot \tilde{q}^q_y(\varrho)\int \mathbb{1}(\psi_{xy}(\varepsilon, G_x^{-1}(\rho), Q_y^{-1}(\varrho)) \geq 0)\mathrm{d}F_{xy}(\varepsilon)\\ &= \alpha u_x v_y \cdot \tilde{g}^q_x(\rho)\cdot \tilde{q}^q_y(\varrho)[1 - F_{xy}(-\phi(\rho,\varrho))]\end{aligned}$$

where $\phi_{xy}(\rho,\varrho) = f_{xy} + G_x^{-1}(\rho) + Q_y^{-1}(\varrho) - rU_x(G_x^{-1}(\rho)) - rV_y(Q_y^{-1}(\varrho))$. Therefore, we know ϕ_{xy} satisfies

$$\begin{cases}\phi_{xy}(\rho,\varrho) = -F_{xy}^{-1}\left(1 - \frac{\delta\mu_{xy}(\rho,\zeta)}{\alpha u_x v_y \tilde{g}^q_x(\rho)\tilde{q}^q_y(\varrho)}\right) & \text{if } \mu_{xy}(\rho,\varrho) > 0\\ \phi_{xy}(\rho,\varrho) \leq -\bar{\varepsilon}_{xy} & \text{if } \mu_{xy}(\rho,\varrho) = 0\end{cases}$$

where $\bar{\varepsilon}_{xy} = \sup\mathrm{supp}(F_{xy})$. Note, the second and third conditions in Assumption 4 imply for any (x,ρ) with $\rho > 0$, there exists some (y',ϱ') such that $\mu_{xy'}(\rho,\varrho') > 0$; for any (y,ϱ) with $\varrho > 0$, there exists some (x',ρ') such that $\mu_{x'y}(\rho',\varrho) > 0$.

Now, $rU_x(G_x^{-1}(\cdot))$ and $rV_y(Q_y^{-1}(\cdot))$ can be calculated, as equation (6) can be rewritten as

$$\begin{aligned}rU_x\left(G_x^{-1}(\rho)\right) &= b_x + \frac{\alpha\theta}{r+\delta}\sum_y v_y \iint \max\left(\varepsilon + \phi_{xy}(\rho,\varrho), 0\right)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}^q_y\left(\varrho\right)\\ rV_y\left(Q_y^{-1}(\varrho)\right) &= c_y + \frac{\alpha\left(1-\theta\right)}{r+\delta}\sum_x u_x \iint \max\left(\varepsilon + \phi_{xy}\left(\rho,\varrho\right), 0\right)\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}\tilde{G}^q_x\left(\rho\right).\end{aligned}$$

Finally, G_x^{-1} and Q_y^{-1}, and, hence, G_x and Q_y are identified as

$$\begin{aligned}G_x^{-1}(\rho) &= \phi_{xy}(\rho,\varrho) - rV_y(Q_y^{-1}(\varrho)) - [\phi_{xy}(0.5, 0.5) - rU_x(G_x^{-1}(0.5)) - rV_y(Q_y^{-1}(0.5))]\\ Q_y^{-1}(\varrho) &= \phi_{xy}(\rho,\varrho) - rU_x(G_x^{-1}(\rho)) - [\phi_{xy}(0.5, 0.5) - rU_x(G_x^{-1}(0.5)) - rV_y(Q_y^{-1}(0.5))]\end{aligned}$$

where G_x and Q_y is normalized to have 0 as median. Therefore, we have proved the following result.

Proposition 4. *Suppose Assumption 1, 2 and 4 hold. Then, the unobserved characteristics distribution* $\{G_x, Q_y : x \in X, y \in Y\}$ *can be identified from the panel data* $\{m_{ijt}, x_i, y_j : i \in I, j \in J, t \in [0,T]\}$ *as* $T \to \infty$, , *given* $\{F_{xy} : x \in X, y \in Y\}$ *and matching technology* α.

Appendix

A Proof of lemma 1

Proof. Rewrite equation (8) as

$$\begin{aligned} \mathring{g}_x(\eta) &= \log\left(\frac{\delta}{\delta+\alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right)\geq 0\right)\exp(\mathring{q}_y(\zeta))\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}Q_y\left(\zeta\right)}\right) \\ \mathring{q}_y(\zeta) &= \log\left(\frac{\delta}{\delta+\alpha\sum_x \mu_x \iint \mathbb{1}\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right)\geq 0\right)\exp(\mathring{g}_x(\eta)\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}G_x\left(\eta\right)}\right) \end{aligned} \tag{29}$$

where

$$\mathring{g}_x(\eta)=\log\left(\frac{u_x\tilde{g}_x(\eta)}{\mu_x g_x(\eta)}\right),\quad \mathring{q}_y(\zeta)=\log\left(\frac{v_y\tilde{q}_y(\zeta)}{\mu_y q_y(\zeta)}\right)$$

Let's now define the space for $\mathring{g}_x$ and $\mathring{q}_y$. Define set Ξ_x and Ξ_y as

$$\Xi_x\equiv\left[\log\left(\frac{\delta}{\delta+\alpha\sum_y\mu_y}\right),0\right]^{\mathrm{supp}(g_x)},\quad \Xi_y\equiv\left[\log\left(\frac{\delta}{\delta+\alpha\sum_x\mu_x}\right),0\right]^{\mathrm{supp}(q_y)}$$

and define set Ξ as

$$\Xi\equiv\prod_{x\in X}\Xi_x\times\prod_{y\in Y}\Xi_y,$$

Next, define subset Ξ_x^c and Ξ_y^c as

$$\begin{aligned} \Xi_x^c &= \{\mathring{g}_x\in\Xi_x:\ \mathring{g}_x \text{ is continuous}\} \\ \Xi_y^c &= \{\mathring{q}_y\in\Xi_y:\ \mathring{q}_y \text{ is continuous}\} \end{aligned}$$

and define $\Xi^c=\prod_x\Xi_x^c\times\prod_y\Xi_y^c$.

Now, let's rewrite equation (29) to self-mappings. Define $\Gamma_x:\Xi^c\mapsto\Xi_x^c$ and $\Gamma_y:\Xi^c\mapsto\Xi_y^c$ as

$$\begin{aligned} \Gamma_x(\{\mathring{g}_x,\mathring{q}_y\})(\eta) &= \log\left(\frac{\delta}{\delta+\alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right)\geq 0\right)e^{\mathring{q}_y(\zeta)}\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}Q_y\left(\zeta\right)}\right) \\ \Gamma_y(\{\mathring{g}_x,\mathring{q}_y\})(\zeta) &= \log\left(\frac{\delta}{\delta+\alpha\sum_x \mu_x \iint \mathbb{1}\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right)\geq 0\right)e^{\mathring{g}_x(\eta)}\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}G_x\left(\eta\right)}\right) \end{aligned}$$

and define $\Gamma:\Xi^c\mapsto\Xi^c$ as $\Gamma(\{\mathring{g}_x,\mathring{q}_y\})=\{\mathring{g}_x',\mathring{q}_y'\}$ where $\mathring{g}_x'=\Gamma_x(\{\mathring{g}_x,\mathring{q}_y\})$ and $\mathring{q}_y'=$

$\Gamma_y(\{\mathring{g}_x, \mathring{q}_y\})$.

For any $\{\mathring{g}_x, \mathring{q}_y\}, \{\mathring{g}'_x, \mathring{q}'_y\} \in \Xi^c$, define metric $d\left(\{\mathring{g}_x, \mathring{q}_y\}, \{\mathring{g}'_x, \mathring{q}'_y\}\right)$ as

$$d\left(\{\mathring{g}_x, \mathring{q}_y\}, \{\mathring{g}'_x, \mathring{q}'_y\}\right) = \max\left(\max_{x\in X}\left(d_x\left(\mathring{g}_x, \mathring{g}'_x\right)\right), \max_{y\in Y}\left(d_y\left(\mathring{q}_y, \mathring{q}'_y\right)\right)\right) \tag{30}$$

where

$$\begin{aligned} d_x\left(\mathring{g}_x, \mathring{g}'_x\right) &= \sup_{\eta\in\operatorname{supp}(g_x)} \left|\mathring{g}_x(\eta) - \mathring{g}'_x(\eta)\right| \\ d_y\left(\mathring{q}_y, \mathring{q}'_y\right) &= \sup_{\zeta\in\operatorname{supp}(q_y)} \left|\mathring{q}_y(\zeta) - \mathring{q}'_y(\zeta)\right|. \end{aligned}$$

Since the set of bounded continuous function with sup norm is a complete metric space (see, theorem 3.1 in Stokey et al. (1989)), Ξ^c is also a complete metric space with metric $d(\cdot,\cdot)$. If T is a contraction mapping in Ξ^c, Γ has a unique fixed point.

To show Γ is a contraction mapping. For any $\{\mathring{g}_x, \mathring{q}_y\}, \{\mathring{g}'_x, \mathring{q}'_y\} \in \Xi^c$, let $\Delta = d\left(\{\mathring{g}_x, \mathring{q}_y\}, \{\mathring{g}'_x, \mathring{q}'_y\}\right)$ and note that for all $x \in X$

$$\begin{aligned} &\left(\Gamma_x\left(\{\mathring{g}_x, \mathring{q}_y\}\right) - \Gamma_x\left(\{\mathring{g}'_x, \mathring{q}'_y\}\right)\right)(\eta) \\ =&\log\left(\frac{\delta + \alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right) e^{\mathring{q}'_y(\zeta)}\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}Q_y(\zeta)}{\delta + \alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right) e^{\mathring{q}_y(\zeta)}\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}Q_y(\zeta)}\right) \\ \leq&\log\left(\frac{\delta + e^{\Delta}\alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right) e^{\mathring{q}_y(\zeta)}\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}Q_y(\zeta)}{\delta + \alpha\sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon,\eta,\zeta)\geq 0\right) e^{\mathring{q}_y(\zeta)}\mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}Q_y(\zeta)}\right) \\ \leq&\log\left(a + (1-a)e^{\Delta}\right) \end{aligned}$$

where

$$a = \frac{\delta}{\delta + \alpha\sum_y \mu_y}.$$

Since $\Delta \in \left[0, \max\left(-\log\left(\frac{\delta}{\delta+\alpha\sum_y\mu_y}\right), -\log\left(\frac{\delta}{\delta+\alpha\sum_x\mu_x}\right)\right)\right]$, by Lemma 3, there exists some $\beta_1 \in (a,1)$ such that $\log\left(a + (1-a)e^{\Delta}\right) \leq \beta_1\Delta$. Note β_1 can be chosen to be independent with $\{rU_x, rV_y, u_x, v_y\}$, as long as $\alpha = m(u,v)/(uv)$ has an upper bound for all u and v of interest. Therefore, we conclude for all $x \in X$,

$$\sup_\eta \left|\left(\Gamma_x\left(\{\mathring{g}_x, \mathring{q}_y\}\right) - \Gamma_x\left(\{\mathring{g}'_x, \mathring{q}'_y\}\right)\right)(\eta)\right| \leq \beta_1\Delta.$$

Similarly, we can show for all $y \in Y$,

$$\sup_{\zeta} \left| \left(\Gamma_y \left(\{ \mathring{g}_x, \mathring{q}_y \} \right) - \Gamma_y \left(\left\{ \mathring{g}'_x, \mathring{q}'_y \right\} \right) \right) (\zeta) \right| \leq \beta_2 \Delta.$$

for some $\beta_2 \in \left(\frac{\delta}{\delta + \alpha \sum_x \mu_x}, 1 \right)$. Similarly, β_2 can be chosen to be independent with $\{rU_x, rV_y, u_x, v_y\}$.

Therefore, $d \left(\Gamma \left(\{ \mathring{g}_x, \mathring{q}_y \} \right) - \Gamma \left(\left\{ \mathring{g}'_x, \mathring{q}'_y \right\} \right) \right) \leq \beta d \left(\{ \mathring{g}_x, \mathring{q}_y \}, \left\{ \mathring{g}'_x, \mathring{q}'_y \right\} \right)$ where $\beta = \max(\beta_1, \beta_2)$, so that Γ is a contraction mapping. Thus, Γ has a unique fixed point $\{\mathring{g}_x, \mathring{q}_y\}$, from which we can get $\{\tilde{g}_x, \tilde{q}_y\}$ as

$$\tilde{g}_x(\eta) = \begin{cases} 0 & \text{if } \eta \notin \text{supp}(g_x), \\ \frac{\mu_x}{u_x} g_x(\eta) \cdot e^{\mathring{g}_x(\eta)} & \text{if } \eta \in \text{supp}(g_x). \end{cases}$$

$$\tilde{q}_y(\zeta) = \begin{cases} 0 & \text{if } \eta \notin \text{supp}(q_y), \\ \frac{\mu_y}{v_y} q_y(\zeta) \cdot e^{\mathring{q}_y(\zeta)} & \text{if } \eta \in \text{supp}(q_y). \end{cases}$$

from which it's easy to see $\{\tilde{g}_x, \tilde{q}_y\}$ are bounded and continuous functions. □

Lemma 3. *Let* $f(x) = \log(a + bx) / \log(x)$ *with* $x \in (1, \infty), a > 0, b > 0$ *and* $a + b = 1$. *Then,* $f(x)$ *is strictly increasing,* $\lim_{x \to 1+} f(x) = b$ *and* $\lim_{x \to \infty} f(x) = 1$.

Proof. Using L'Hôpital's rule, one can easy derive $\lim_{x \to 1+} f(x) = b$ and $\lim_{x \to \infty} f(x) = 1$. To show $f(x)$ is strictly increasing, take derivative of $f(x)$ as

$$f'(x) = \frac{bx \log(x) - (a + bx) \log(a + bx)}{(a + bx) \cdot x \cdot (\log x)^2}.$$

let $g(x) = bx \cdot \log(x)$, $h(x) = (a + bx) \log(a + bx)$. $f'(x) > 0$ if and only if $g(x) > h(x)$.

Note $g(1) = h(1) = 0$, and $g'(x) > h'(x)$ for all $x > 1$. Therefore, $g(x) > h(x)$ for all $x > 1$. Hence, $f(x)$ is strictly increasing in x. □

B Existence of Stationary Equilibrium

In this section, we prove the existence of a SE under a slightly general setting without imposing Assumption 1.

For man i with characteristics (x, η) and woman j with (y, ζ), their matching surplus $f(i, j)$ equals to $f_{xy}(\eta, \zeta) + \varepsilon$ instead of $f_{xy} + \eta + \zeta + \varepsilon$, where $f_{xy}(\cdot, \cdot)$ is a group-pairwise matching surplus function on unobserved characteristics, and ε is a matching specific random shock distributed as F_{xy}. Therefore, ψ_{xy} in equation (7) can now be rewritten as

$$\psi_{xy}\left(\varepsilon,\eta,\zeta\right)\equiv f_{xy}(\eta,\zeta)+\varepsilon-rU_{x}\left(\eta\right)-rV_{y}\left(\zeta\right) \tag{31}$$

Assume $f_{xy}(\eta,\zeta)$ is a differentiable function in η and ζ. Moreover, we assume η and ζ are in a compact set, and matching function $m(u,v)$ is either quadratic or Cobb–Douglas. We collect all regularity conditions in the following assumption.

Assumption 5 (Regularity Conditions for SE). *There exists some $\bar{d}>0$, such that for each $x\in X$ and $y\in Y$,*

(i) $f_{xy}(\eta,\zeta)$ is differentiable in η and ζ, and for all $\eta\in\operatorname{supp}(G_x)$ and $\zeta\in\operatorname{supp}(Q_y)$,

$$0\leq\frac{\partial f_{xy}(\eta,\zeta)}{\partial\eta}\leq\bar{d},\quad 0\leq\frac{\partial f_{xy}(\eta,\zeta)}{\partial\zeta}\leq\bar{d},$$

(ii) $\operatorname{supp}(G_x)$ and $\operatorname{supp}(Q_y)$ are compact,

(iii) F_{xy} is absolutely continuous, and its density is bounded in $[0,\bar{d}]$.

Moreover, the matching technology $m(u,v)$ takes one of the following forms.

(a) $m(u,v)=\alpha uv$ with constant α,

(b) $m(u,v)=\gamma u^{\kappa}v^{1-\kappa}$ for some $\gamma>0$ and $\kappa\in(0,1)$.

Finally, without loss of generality, we normalize $b=c=0$. Therefore, the equilibrium condition (6) can be written as

$$\begin{aligned}
rU_{x}\left(\eta\right)&=\frac{\alpha\theta}{r+\delta}\sum_{y}v_{y}\iint\max\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right),0\right)\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}\tilde{Q}_{y}\left(\zeta\right)\\
rV_{y}\left(\zeta\right)&=\frac{\alpha\left(1-\theta\right)}{r+\delta}\sum_{x}u_{x}\iint\max\left(\psi_{xy}\left(\varepsilon,\eta,\zeta\right),0\right)\mathrm{d}F_{xy}\left(\varepsilon\right)\mathrm{d}\tilde{G}_{x}\left(\eta\right)
\end{aligned} \tag{32}$$

Proposition 5 (Existence of SE). *Under Assumption 2 and 5, there exists a stationary equilibrium $\{U_x,V_y,\tilde{G}_x,\tilde{Q}_y,u_x,v_y : x\in X,y\in Y\}$ satisfying condition (32), (31), (8) and (9).*

Proof. We first introduce some useful notation. For any $x\in X$ and $y\in Y$, define $\mathcal{C}_x$ and $\mathcal{C}_y$

as

$$\mathcal{C}_x = \left\{U_x \in \mathbb{R}^{\text{supp}(G_x)} : U_x \text{ is continuous function in } \text{supp}(G_x)\right\},$$
$$\mathcal{C}_y = \left\{V_y \in \mathbb{R}^{\text{supp}(Q_y)} : V_y \text{ is continuous function in } \text{supp}(Q_y)\right\}.$$

And define $\mathcal{C} = \prod_x \mathcal{C}_x \times \prod_x \mathcal{C}_y$. Define the norm $\|\cdot\|_{\mathscr{C}}$ on $\mathscr{C} \equiv \mathbb{R}^{|X|+|Y|} \times \mathcal{C}$ as

$$\|\{u_x, v_y, U_x, V_y\}\|_{\mathscr{C}} = \max\Big(\max_{x\in X}\left(|u_x|, \text{supp}_\eta |U_x(\eta)|\right) \quad \max_{y\in Y}\left(|v_y|, \text{supp}_\zeta |V_y(\zeta)|\right)\Big) \tag{33}$$

Since $\mathbb{R}$, $\mathcal{C}_x$ and $\mathcal{C}_y$ are all Banach space with supremum norms, $\mathscr{C}$ is also a Banach space with norm $\|\cdot\|_{\mathscr{C}}$.

Next, define $\mathcal{D}_x$ and $\mathcal{D}_y$ as

$$\mathcal{D}_x = \left\{U_x \in \mathcal{C}_x : U_x \text{ is differentiable and } \forall \eta \in \text{supp}(G_x), 0 \leq \frac{\mathrm{d}rU_x(\eta)}{\mathrm{d}\eta} \leq \bar{d}.\right\},$$
$$\mathcal{D}_y = \left\{V_y \in \mathcal{C}_y : V_y \text{ is differentiable and } \forall \zeta \in \text{supp}(Q_y), 0 \leq \frac{\mathrm{d}rV_y(\zeta)}{\mathrm{d}\zeta} \leq \bar{d}.\right\}.$$

And, let $\mathcal{D} = \prod_x \mathcal{D}_x \times \prod_x \mathcal{D}_y$.

We divide the remainder of the proof into two parts for two types of matching technology $m(\cdot,\cdot)$ respectively.

Suppose Condition (*a*) in Assumption 5 hold. Define $t, s \in (0,1)$ as

$$t = \frac{1}{1 + \frac{\alpha\theta}{r+\delta}\sum_y \mu_y}$$
$$s = \frac{1}{1 + \frac{\alpha(1-\theta)}{r+\delta}\sum_x \mu_x}$$

For each $x \in X$ and $y \in Y$, define $T_x : \mathscr{C} \mapsto \mathcal{C}_x$ and $T_y : \mathscr{C} \mapsto \mathcal{C}_y$ as

$$T_x(\{u_x, v_y, U_x, V_y\})(\eta) = t \cdot \frac{\alpha\theta}{r(r+\delta)} \sum_y v_y \iint \max\left(\psi_{xy}\left(\varepsilon, \eta, \zeta\right), 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{Q}_y\left(\zeta\right) + (1-t)U_x(\eta)$$
$$T_y(\{u_x, v_y, U_x, V_y\})(\zeta) = s \cdot \frac{\alpha\left(1-\theta\right)}{r(r+\delta)} \sum_x u_x \iint \max\left(\psi_{xy}\left(\varepsilon, \eta, \zeta\right), 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{G}_x\left(\eta\right) + (1-s)V_y(\zeta)$$

where $\{U_x, V_y\}$ enters into $\{\psi_{xy}\}$ through (31) and $\tilde{G}_x$ and $\tilde{Q}_y$ is pinned down by $\{u_x, v_y, U_x, V_y\}$ as in lemma 1.

Define $H_x : \mathscr{C} \mapsto \mathbb{R}$ and $H_y : \mathscr{C} \mapsto \mathbb{R}$ as

$$\begin{aligned} H_x(\{u_x, v_y, U_x, V_y\}) &= \frac{\delta \mu_x}{\delta + \alpha \sum_y v_y \iiint \mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0) \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}\tilde{G}_x(\eta) \mathrm{d}\tilde{Q}_y(\zeta)} \\ H_y(\{u_x, v_y, U_x, V_y\}) &= \frac{\delta \mu_y}{\delta + \alpha \sum_x u_x \iiint \mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0) \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}\tilde{G}_x(\eta) \mathrm{d}\tilde{Q}_y(\zeta)} \end{aligned}$$

Define $\Upsilon : \mathscr{C} \mapsto \mathscr{C}$ as $\Upsilon(\{u_x, v_y, U_x, V_y\}) = \left\{u'_x, v'_y, U'_x, V'_y\right\}$ where

$$\begin{aligned} u'_x &= H_x(\{u_x, v_y, U_x, V_y\}) \\ v'_y &= H_y(\{u_x, v_y, U_x, V_y\}) \\ U'_x &= T_x(\{u_x, v_y, U_x, V_y\}) \\ V'_y &= T_y(\{u_x, v_y, U_x, V_y\}) \end{aligned}$$

It's easy to see that a stationary equilibrium exists if and only if Υ has a fixed point. Moreover, Υ is a continuous mapping on $\mathscr{C}$, by lemma 5.

Define K_1, K_2 and K as

$$\begin{aligned} K_1 \equiv & \prod_{x \in X} [0, \mu_x] \times \prod_{y \in Y} [0, \mu_y] \\ K_2 \equiv & \Big\{ \{U_x, V_y\} \in \mathcal{D} : \forall x \in X, y \in Y, \eta \in \mathrm{supp}(G_x), \zeta \in \mathrm{supp}(Q_y), \\ & 0 \leq r U_x(\eta) \leq \frac{\alpha \theta}{r + \delta} \sum_y \mu_y \iint |f_{xy}(\eta, \zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}Q_y(\zeta) \\ & 0 \leq r V_y(\zeta) \leq \frac{\alpha(1-\theta)}{r + \delta} \sum_x \mu_x \iint |f_{xy}(\eta, \zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}G_y(\eta) \Big\} \\ K \equiv & K_1 \times K_2 \end{aligned} \tag{34}$$

By lemma 4, K is a totally bounded subset in $\mathscr{C}$. Hence, the its closure cl$K$ in $\mathscr{C}$ is compact.

We first check $\Upsilon(K) \subset K$. For any $\{u_x, v_y, U_x, V_y\} \in K$, it's trivial to see $H_x(\{u_x, v_y, U_x, V_y\}) \in [0, \mu_x]$ and $H_y(\{u_x, v_y, U_x, V_y\}) \in [0, \mu_y]$. $T_x(\{u_x, v_y, U_x, V_y\})$ is differentiable under Condition (i) in Assumption 5. Moreover, for $\left\{u'_x, v'_y, U'_x, V'_y\right\} = \Upsilon(\{u_x, v_y, U_x, V_y\})$,

$$
\begin{aligned}
U_x'(\eta) \le & t \cdot \frac{\alpha\theta}{(r+\delta)} \sum_y \mu_y \iint \max\left(f_{xy}(\eta,\zeta)+\varepsilon, 0\right) \mathrm{d}F_{xy}(\varepsilon)\,\mathrm{d}Q_y(\zeta) \\
& + (1-t)U_x(\eta) \\
\le & \frac{\alpha\theta}{r+\delta} \sum_y \mu_y \iint \left|f_{xy}(\eta,\zeta)+\varepsilon\right| \mathrm{d}F_{xy}(\varepsilon)\mathrm{d}Q_y(\zeta)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\mathrm{d}rU_x'(\eta)}{\mathrm{d}\eta} =& t \cdot \frac{\alpha\theta}{r+\delta} \sum_y v_y \iint \mathbb{1}(\psi(\varepsilon,\eta,\zeta)\ge 0) \frac{\partial f_{xy}(\eta,\zeta)}{\partial\eta} \mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}_y(\zeta) \\
& + \frac{\mathrm{d}rU_x(\eta)}{\mathrm{d}\eta}\left((1-t) - t\cdot\frac{\alpha\theta}{r+\delta}\sum_y v_y \iint \mathbb{1}(\psi(\varepsilon,\eta,\zeta)\ge 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}_y(\zeta)\right) \\
\le & t\bar{d}\cdot\frac{\alpha\theta}{r+\delta}\sum_y v_y \iint \mathbb{1}(\psi(\varepsilon,\eta,\zeta)\ge 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}_y(\zeta)\cdot \\
& + \bar{d}\left((1-t) - t\cdot\frac{\alpha\theta}{r+\delta}\sum_y v_y \iint \mathbb{1}(\psi(\varepsilon,\eta,\zeta)\ge 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}_y(\zeta)\right) \qquad (35) \\
\le & \bar{d}
\end{aligned}
$$

where inequality (35) comes from the fact that

$$
\begin{aligned}
& (1-t) - t\cdot\frac{\alpha\theta}{r+\delta}\sum_y v_y \iint \mathbb{1}(\psi(\varepsilon,\eta,\zeta)\ge 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{Q}_y(\zeta) \\
\ge & (1-t) - t\cdot\frac{\alpha\theta}{r+\delta}\sum_y \mu_y \\
= & 0
\end{aligned}
$$

Similarly, one can show

$$
\begin{aligned}
rV_x'(\zeta) \le & \frac{\alpha(1-\theta)}{r+\delta}\sum_x \mu_x \iint \left|f_{xy}(\eta,\zeta)+\varepsilon\right| \mathrm{d}F_{xy}(\varepsilon)\mathrm{d}G_y(\eta)\Big\} \\
\frac{\mathrm{d}rU_x'(\eta)}{\mathrm{d}\eta} \le & \bar{d}
\end{aligned}
$$

Therefore, $\Upsilon(K) \subset K$.

By lemma 5, Υ is a continuous mapping on $\mathscr{C}$. Therefore, $\Upsilon(\mathrm{cl}K) \subset \mathrm{cl}K$. Moreover, K

is a convex set by construction, so that cl$K$ is also convex. By Schauder-Tychonoff fixed point theorem[4], Υ has a fixed point in cl$K$.

Suppose Condition (*b*) in Assumption 5 hold. Define $\hat{K}_1 \equiv \prod_{x\in X} \left[\underline{u_x}, \mu_x\right] \times \prod_{y\in Y} \left[\underline{v_y}, \mu_y\right]$, where $\left\{\underline{u_x}, \underline{v_y}\right\}$ solves

$$\delta \underline{u_x} + \gamma \cdot \underline{u_x}^{\kappa} \left(\sum_y \mu_y\right)^{1-\kappa} = \delta \mu_x$$
$$\delta \underline{v_y} + \gamma \cdot \left(\sum_x \mu_x\right)^{\kappa} \underline{v_y}^{1-\kappa} = \delta \mu_y$$

Such $\left\{\underline{u_x}, \underline{v_y}\right\}$ always exists and is unique, and for all x and y, $\underline{u_x} > 0$ and $\underline{v_y} > 0$. Therefore, for any $\{u_x, v_y\} \in \hat{K}_1$, $\underline{\alpha} \leq \frac{m(u,v)}{uv} \leq \bar{\alpha}$ where

$$\underline{\alpha} = \gamma \left(\sum_x \mu_x\right)^{\kappa-1} \left(\sum_x \mu_y\right)^{-\kappa}, \ \bar{\alpha} = \gamma \left(\sum_x \underline{u_x}\right)^{\kappa-1} \left(\sum_y \underline{v_y}\right)^{-\kappa}.$$

Define $\hat{K}_2$, $\hat{K}$ and $\hat{\mathscr{C}}$ as

$$\begin{aligned}
\hat{K}_2 \equiv & \Big\{ \{U_x, V_y\} \in \mathcal{D} : \forall x \in X, y \in Y, \eta \in \operatorname{supp}(G_x), \zeta \in \operatorname{supp}(Q_y), \\
& 0 \leq rU_x(\eta) \leq \frac{\bar{\alpha}\theta}{r+\delta} \sum_y \mu_y \iint |f_{xy}(\eta,\zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}Q_y(\zeta) \\
& 0 \leq rV_y(\zeta) \leq \frac{\bar{\alpha}(1-\theta)}{r+\delta} \sum_x \mu_x \iint |f_{xy}(\eta,\zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}G_x(\eta) \Big\} \\
\hat{K} \equiv & \hat{K}_1 \times \hat{K}_2 \\
\hat{\mathscr{C}} \equiv & \hat{K}_1 \times \mathcal{C}
\end{aligned} \tag{36}$$

Note $\hat{\mathscr{C}}$ is a complete local convex space, when $\hat{\mathscr{C}}$ is viewed as a subset of $\mathscr{C}$.

Define $\hat{t}, \hat{s} \in (0,1)$ as

$$\hat{t} = \frac{1}{1 + \frac{\bar{\alpha}\theta}{r+\delta} \sum_y \mu_y}$$
$$\hat{s} = \frac{1}{1 + \frac{\bar{\alpha}(1-\theta)}{r+\delta} \sum_x \mu_x}$$

[4] See, for example, Theorem 5.28 in Rudin (1991) for a reference.

For each $x \in X$ and $y \in Y$, define $\hat{T}_x : \hat{\mathscr{C}} \mapsto \mathcal{C}_x$ and $\hat{T}_y : \hat{\mathscr{C}} \mapsto \mathcal{C}_y$ as

$$\hat{T}_x(\{u_x, v_y, U_x, V_y\})(\eta) = \hat{t} \cdot \frac{\alpha\theta}{r(r+\delta)} \sum_y v_y \iint \max\left(\psi_{xy}\left(\varepsilon, \eta, \zeta\right), 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{Q}_y\left(\zeta\right) \\ + (1-\hat{t})U_x(\eta)$$

$$\hat{T}_y(\{u_x, v_y, U_x, V_y\})(\zeta) = \hat{s} \cdot \frac{\alpha\left(1-\theta\right)}{r(r+\delta)} \sum_x u_x \iint \max\left(\psi_{xy}\left(\varepsilon, \eta, \zeta\right), 0\right) \mathrm{d}F_{xy}\left(\varepsilon\right) \mathrm{d}\tilde{G}_x\left(\eta\right) \\ + (1-\hat{s})V_y(\zeta)$$

where α here should be understood as $m(u, v)/(uv)$, $\{U_x, V_y\}$ enters into $\{\psi_{xy}\}$ by (31)and $\tilde{G}_x$ and $\tilde{Q}_y$ is pinned down by $\{u_x, v_y, U_x, V_y\}$ as in lemma 1.

Define $\hat{H}_x : \hat{\mathscr{C}} \mapsto \mathbb{R}$ and $\hat{H}_y : \hat{\mathscr{C}} \mapsto \mathbb{R}$ as

$$\hat{H}_x(\{u_x, v_y, U_x, V_y\}) = \frac{\delta\mu_x}{\delta + \alpha\sum_y v_y \iiint \mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{G}_x(\eta)\mathrm{d}\tilde{Q}_y(\zeta)}$$
$$\hat{H}_y(\{u_x, v_y, U_x, V_y\}) = \frac{\delta\mu_y}{\delta + \alpha\sum_x u_x \iiint \mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0)\mathrm{d}F_{xy}(\varepsilon)\mathrm{d}\tilde{G}_x(\eta)\mathrm{d}\tilde{Q}_y(\zeta)}$$

Define $\hat{\Upsilon} : \hat{\mathscr{C}} \mapsto \mathscr{C}$ as $\hat{\Upsilon}(\{u_x, v_y, U_x, V_y\}) = \left\{u'_x, v'_y, U'_x, V'_y\right\}$ where

$$u'_x = \hat{H}_x(\{u_x, v_y, U_x, V_y\})$$
$$v'_y = \hat{H}_y(\{u_x, v_y, U_x, V_y\})$$
$$U'_x = \hat{T}_x(\{u_x, v_y, U_x, V_y\})$$
$$V'_y = \hat{T}_y(\{u_x, v_y, U_x, V_y\})$$

It's easy to see that a stationary equilibrium exists if and only if $\hat{\Upsilon}$ has a fixed point. Moreover, $\hat{\Upsilon}$ is a continuous mapping on $\hat{\mathscr{C}}$, by lemma 5.

We first check $\hat{\Upsilon}(\hat{K}) \subset \hat{K}$. For any $\{u_x, v_y, U_x, V_y\} \in \hat{K}$, let $\left\{u'_x, v'_y, U'_x, V'_y\right\} = \hat{\Upsilon}(\{u_x, v_y, U_x, V_y\})$.

To see $u'_x \in [\underline{u_x}, \mu_x]$, note that

$$\begin{aligned} u'_x &\geq \frac{\delta \mu_x}{\delta + \alpha \sum_y v_y} \\ &= \frac{\delta \mu_x}{\delta + \gamma \left(\frac{v}{u}\right)^{1-\kappa}} \\ &\geq \frac{\delta \mu_x}{\delta + \gamma \left(\frac{\sum_y \mu_y}{\sum_x \underline{u_x}}\right)^{1-\kappa}} \\ &= \underline{u_x}. \end{aligned}$$

Similarly, we can show $v'_y \in [\underline{v_y}, \mu_y]$. Similar to the proof for $\Upsilon(K) \subset K$, one can show $\{U'_x, V'_y\} \in \hat{K}_2$. Hence, $\hat{\Upsilon}(\hat{K}) \subset \hat{K}$.

By lemma 5, $\hat{\Upsilon}$ is a continuous mapping on $\mathscr{C}$. Therefore, $\hat{\Upsilon}(\text{cl}\hat{K}) \subset \text{cl}\hat{K}$. Moreover, $\hat{K}$ is a convex set by construction, so that $\text{cl}\hat{K}$ is also convex. By Schauder-Tychonoff fixed point theorem, $\hat{\Upsilon}$ has a fixed point in $\text{cl}\hat{K}$.

□

Lemma 4. *Under Assumption 2, and 5, the set K defined in (34) is totally bounded in $\mathscr{C}$, and the set $\hat{K}$ defined in (36) is totally bounded in $\hat{\mathscr{C}}$.*

Proof. Since the proof for $\hat{K}$ is almost the same as K, we only focus on K in this proof.

For any $x \in X$, $[0, \mu_x]$ is a compact set in $\mathbb{R}$. Let $K_{2,x}$ be defined as

$$\begin{aligned} K_{2,x} \equiv \Big\{ & U_x \in \mathcal{D}_x : \forall \eta \in \text{supp}(G_x), \\ & 0 \leq r U_x(\eta) \leq \frac{\alpha\theta}{r+\delta} \sum_y \mu_y \iint |f_{xy}(\eta,\zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}Q_y(\zeta). \Big\} \end{aligned}$$

We establish two facts about $K_{2,x}$.

(i) By construction, for any $U_x \in K_{2,x}$,

$$\sup_{\eta \in \text{supp}(G_x)} |U_x(\eta)| \leq \sup_{\eta \in \text{supp}(G_x)} \frac{\alpha\theta}{r+\delta} \sum_y \mu_y \iint |f_{xy}(\eta,\zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon) \mathrm{d}Q_y(\zeta).$$

Thus, $K_{2,x}$ is bounded in supremum norm.

(ii) For any $\eta \in \text{supp}(G_x)$ and $\epsilon > 0$, set $\rho = r\epsilon/\bar{d}$. Then, for any η' with $|\eta' - \eta| < \rho$ and for any $U_x \in K_{2,x}$,

$$|U_x(\eta) - U_x(\eta')| \leq \frac{\bar{d}}{r}|\eta - \eta'| < \epsilon.$$

Thus, $K_{2,x}$ is equicontinuous.

By Arzelà–Ascoli theorem[5], these two facts imply $K_{2,x}$ is totally bounded in supremum norm, and cl$K_{2,x}$ is compact. Similarly, we can show $K_{2,y}$ is totally bounded for any $y \in Y$, where

$$K_{2,y} \equiv \Big\{ V_y \in \mathcal{D}_y : \forall \zeta \in \text{supp}(Q_y), \\ 0 \leq rV_y(\zeta) \leq \frac{\alpha(1-\theta)}{r+\delta} \sum_x \mu_x \iint |f_{xy}(\eta,\zeta) + \varepsilon| \, \mathrm{d}F_{xy}(\varepsilon)\mathrm{d}G_x(\eta). \Big\}$$

Therefore, $K_2 = \prod_{x \in X} K_{2,x} \times \prod_{y \in Y} K_{2,y}$ is also totally bounded, and clK_2 is compact. Finally, K is totally bounded and cl$K$ is compact in L. □

Lemma 5. *Under Assumption 2 and 5, Υ defined in Proposition 5 are continuous in $\mathscr{C}$, and $\hat{\Upsilon}$ defined in Proposition 5 are continuous in $\hat{\mathscr{C}}$*

Proof. Define norm $\|\cdot\|_{\mathcal{C}} : \mathcal{C} \mapsto \mathbb{R}_+$ as

$$\|\{\tilde{g}_x, \tilde{q}_y\}\|_{\mathcal{C}} = \max\left(\max_{x \in X} \left(\sup_{\eta \in \text{supp}(G_x)} |\tilde{g}(\eta)| \right), \max_{y \in Y} \left(\sup_{\zeta \in \text{supp}(Q_y)} |\tilde{q}(\zeta)| \right) \right)$$

We first show the endogenous conditional density $\{\tilde{g}_x, \tilde{q}_y\}$ as an implicit function of $\{u_x, v_y, U_x, V_y\}$ defined in equation (8) is continuous when $\{u_x, v_y, U_x, V_y\} \in \mathscr{C}$ and Condition (a) in Assumption 5 hold. The topology for $\{\tilde{g}_x, \tilde{q}_y\}$ is induced by norm $\|\cdot\|_{\mathcal{C}}$.

For any $\{u_x, v_y, U_x, V_y\}$ and $\{u'_x, v'_y, U'_x, V'_y\}$ in $\mathscr{C}$, let $\{\tilde{g}_x, \tilde{q}_y\}$ and $\{\tilde{g}'_x, \tilde{q}'_y\}$ denote their associated density of endogenous conditional distribution solved in equation (8). As in the proof of Lemma 1, define

$$\mathring{g}_x(\eta) = \log\left(\frac{u_x \tilde{g}_x(\eta)}{\mu_x g_x(\eta)}\right), \quad \mathring{q}_y(\zeta) = \log\left(\frac{v_y \tilde{q}_y(\zeta)}{\mu_y q_y(\zeta)}\right)$$

Similarly, we can define $\{\mathring{g}'_x, \mathring{q}'_y\}$ with $\{u'_x, v'_y, \tilde{g}'_x, \tilde{q}'_y\}$. Recall the contraction mapping Γ defined in lemma 1 is associated with $\{u_x, v_y, U_x, V_y\}$. Let Γ' denote the one associated with $\{u'_x, v'_y, U'_x, V'_y\}$. Note Γ and Γ' share the same contraction coefficient. Then, the fact that

[5] See Hanche-Olsen and Holden (2010) for a good reference on Arzelà–Ascoli theorem.

Γ' is a contraction mapping implies

$$\left\|\left\{\mathring{g}'_x, \mathring{q}'_y\right\} - \{\mathring{g}_x, \mathring{q}_y\}\right\|_{\mathcal{C}} \leq \frac{1}{1-\beta} \cdot \|\Gamma'(\{\mathring{g}_x, \mathring{q}_y\}) - \{\mathring{g}_x, \mathring{q}_y\}\|_{\mathcal{C}}$$

For any $x \in X$, let ψ'_{xy} be the one associated with $\{u_x, v_y, U_x, V_y\}$ by equation (31)

$$\begin{aligned}
&|\Gamma'_x(\{\mathring{g}_x, \mathring{q}_y\})(\eta) - \mathring{g}_x(\eta)| \\
=&\left|\log\left(\frac{\delta + \alpha \sum_y \mu_y \iint \mathbb{1}\left(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0\right) e^{\mathring{q}_y(\zeta)} \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta)}{\delta + \alpha \sum_y \mu_y \iint \mathbb{1}\left(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0\right) e^{\mathring{q}_y(\zeta)} \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta)}\right)\right| \\
=&\left|\log\left(1 + \frac{\alpha \sum_y \mu_y \iint \left[\mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0) - \mathbb{1}(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0)\right] e^{\mathring{q}_y(\zeta)} \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta)}{\delta + \alpha \sum_y \mu_y \iint \mathbb{1}\left(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0\right) e^{\mathring{q}_y(\zeta)} \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta)}\right)\right|
\end{aligned}$$

Since

$$\delta + \alpha \sum_y \mu_y \iint \mathbb{1}\left(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0\right) e^{\mathring{q}_y(\zeta)} \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta)$$

is bounded between δ and $\delta + \alpha \sum_y \mu_y$, and since $\exp(\mathring{q}_y(\zeta)) \in [0, 1]$, $|\Gamma'_x(\{\mathring{g}_x, \mathring{q}_y\})(\eta) - \mathring{g}_x(\eta)|$ would converge to 0 if the following term converges to 0,

$$\iint \left|\mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0) - \mathbb{1}(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0)\right| \mathrm{d}F_{xy}(\varepsilon)\, \mathrm{d}Q_y(\zeta).$$

This can be done uniformly over all $\eta \in \operatorname{supp}(G_x)$ if the distance of $\{u_x, v_y, U_x, V_y\}$ and $\left\{u'_x, v'_y, U'_x, V'_y\right\}$ converges to 0 in norm $\|\cdot\|_{\mathscr{C}}$. To see this, note for any $\eta \in \operatorname{supp}(G_x)$,

$$\begin{aligned}
&\int \left|\mathbb{1}(\psi_{xy}(\varepsilon, \eta, \zeta) \geq 0) - \mathbb{1}(\psi'_{xy}(\varepsilon, \eta, \zeta) \geq 0)\right| \mathrm{d}F_{xy}(\varepsilon) \\
=&\int \mathbb{1}(-f_{xy}(\eta, \zeta) + rU_x(\eta) + rV_y(\zeta) \leq \varepsilon \leq -f_{xy}(\eta, \zeta) + rU'_x(\eta) + rV'_y(\zeta)) \mathrm{d}F_{xy}(\varepsilon) \\
+&\int \mathbb{1}(-f_{xy}(\eta, \zeta) + rU'_x(\eta) + rV'_y(\zeta) \leq \varepsilon \leq -f_{xy}(\eta, \zeta) + rU_x(\eta) + rV_y(\zeta)) \mathrm{d}F_{xy}(\varepsilon) \\
\leq& 2 \sup_\varepsilon(\mathrm{d}F_{xy}(\varepsilon)) \cdot \left[\sup_\eta(rU_x(\eta) - rU'_x(\eta)) + \sup_\zeta(rV_y(\zeta) - rV'_y(\zeta))\right].
\end{aligned}$$

where the density $\mathrm{d}F_{xy}(\varepsilon)$ of F_{xy} is bounded as assumed in Condition (iii) in Assumption 5.

Similarly, for any $y \in Y$, $\sup_\zeta \left|\Gamma'_y(\{\mathring{g}_x, \mathring{q}_y\})(\zeta) - \mathring{q}_y(\zeta)\right| \to 0$ as the distance of $\{u_x, v_y, U_x, V_y\}$ and $\left\{u'_x, v'_y, U'_x, V'_y\right\}$ converges to 0.

Therefore, $\{\tilde{g}_x, \tilde{q}_y\}$ as an implicit function of $\{u_x, v_y, U_x, V_y\}$ defined in equation (8) is continuous when $\{u_x, v_y, U_x, V_y\} \in \mathscr{C}$ and Condition (a) in Assumption 5 hold.

Next, we show the endogenous conditional density $\{\tilde{g}_x, \tilde{q}_y\}$ as an implicit function of $\{u_x, v_y, U_x, V_y\}$ defined in equation (8) is continuous when $\{u_x, v_y, U_x, V_y\} \in \hat{\mathscr{C}}$ and Condition (b) in Assumption 5 hold. The topology for $\{\tilde{g}_x, \tilde{q}_y\}$ is again induced by norm $\|\cdot\|_{\mathcal{C}}$.

To show this result, note that $\alpha = m(u, v)/(uv)$ has an upper bound when $\{u_x, v_y\} \in \hat{K}_1$. Therefore, contraction mapping Γ in the proof of Lemma 1 still shares the same contraction coefficient for all $\{u_x, v_y, U_x, V_y\} \in \hat{\mathscr{C}}$. The rest of the proof is the same as the previous case.

Given these two results, the continuity of Υ and $\hat{\Upsilon}$ are now trivial to show.

□

C Proof of lemma 2

Proof. Suppose $\{f_{xy}\}$ satisy the assumption of the lemma. We need to show, for any $x \in X$ and $y \in Y$, $\delta\mu_{xy} < \alpha u_x v_y$. Under Assumption 2, we only need to show there exists some $(\varepsilon, \eta, \zeta) \in \text{int}\left(\text{supp}\left(F_{xy}\right) \times \text{supp}\left(G_x\right) \times \text{supp}\left(Q_y\right)\right)$ such that

$$f_{xy} + \varepsilon + \eta + \zeta - rU_x\left(\eta\right) - rV_y\left(\zeta\right) < 0 \tag{37}$$

Since $rU_x(\eta) \geq b_x$ and $rV_y(\zeta) \geq c_y$, we know

$$f_{xy} + \underline{\varepsilon} + \underline{\eta} + \underline{\zeta} - rU_x\left(\underline{\eta}\right) - rV_y\left(\underline{\zeta}\right) < 0$$

Since $rU_x(\cdot)$ and $rV_y(\cdot)$ are continuous, there exists some neighborhood of $(\underline{\varepsilon}, \underline{\eta}, \underline{\zeta})$, such that any $(\varepsilon, \eta, \zeta)$ in such neighborhood satisfies inequality (37). □

D Proof of Proposition 2

Proof. By equation (21), we only need to show that $\{h_x(\cdot), h_y(\cdot)\}$ can be identified from the distribution of unmatched duration $\{D_x, D_y\}$.

In section 6.1, we have shown that $h_x(\cdot)$ and $h_y(\cdot)$ is non-decreasing. In addition, under Assumption 1, 2 and conditions in Proposition 2, it's easy to check any $\{h_x(\cdot), h_y(\cdot)\}$ defined by (11) and (12) is in set H, where

$$H = \Big\{\{h_x(\cdot), h_y(\cdot)\} \in \prod_{x\in X}[0,\alpha v]^{\text{supp}(G_x)} \times \prod_{y\in Y}[0,\alpha u]^{\text{supp}(Q_y)} : \forall x \in X, \forall y \in Y,$$
$$h_x(\cdot) \text{ and } h_y(\cdot) \text{ are non-decreasing and continuously differentiable;}$$
$$\text{if } h_x(\eta) \in (0,\alpha v), \text{ then } h_x'(\eta) > 0; \text{if } h_y(\zeta) \in (0,\alpha u), \text{ then } h_y'(\zeta) > 0;$$
$$\text{If } \sup\text{supp}(G_x) = \infty, \text{ then } \lim_{\eta\to\infty} h_x(\eta) = \alpha v; \text{If } \sup\text{supp}(Q_y) = \infty,$$
$$\text{then } \lim_{\zeta\to\infty} h_y(\zeta) = \alpha u; \Big\}.$$

In particular, we know from equation (11) that

$$\begin{aligned}
h_x'(\eta) =& \alpha \sum_y v_y \frac{\partial}{\partial \eta} \int 1 - F_{xy}(-f_{xy} - \eta - \zeta + rU_x(\eta) + rV_y(\zeta)) \mathrm{d}\tilde{Q}_y(\zeta) \\
=& \alpha \sum_y v_y \left(1 - \frac{\partial r U_x(\eta)}{\partial \eta}\right) \int dF_{xy}(-f_{xy} - \eta - \zeta + rU_x(\eta) + rV_y(\zeta)) \mathrm{d}\tilde{Q}_y(\zeta) \\
=& \alpha \sum_y \frac{v_y}{1 + \frac{\theta}{r+\delta} h_x(\eta)} \int dF_{xy}(-f_{xy} - \eta - \zeta + rU_x(\eta) + rV_y(\zeta)) \mathrm{d}\tilde{Q}_y(\zeta)
\end{aligned}$$

where $dF_{xy}(\cdot)$ stands for the density function of distribution F_{xy}.

Suppose there exist $\{h_x(\cdot), h_y(\cdot)\}$ and $\{\tilde{h}_x(\cdot), \tilde{h}_y(\cdot)\}$ in H such that, for all x and y, and for any $l > 0$,

$$\begin{aligned}
\int \frac{e^{-l h_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta) \mathrm{d}\eta &= \int \frac{e^{-l \tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)} g_x(\eta) \mathrm{d}\eta \\
\int \frac{e^{-l h_y(\zeta)}}{\delta + h_y(\zeta)} q_y(\zeta) \mathrm{d}\zeta &= \int \frac{e^{-l \tilde{h}_y(\zeta)}}{\delta + \tilde{h}_y(\zeta)} q_y(\zeta) \mathrm{d}\zeta
\end{aligned} \tag{38}$$

We want to show $h_x = \tilde{h}_x$ and $h_y = \tilde{h}_y$ for any group x and y.

For now, fix an arbitrary group x. Let $\underline{\eta} = \inf \text{supp}(G_x)$ and $\bar{\eta} = \sup \text{supp}(G_x)$. Define

$$\begin{aligned}
A &= \{\eta \in \text{supp}(G_x) : h_x(\eta) > \tilde{h}_x(\eta)\}, \\
B &= \{\eta \in \text{supp}(G_x) : h_x(\eta) < \tilde{h}_x(\eta)\}, \\
C &= \{\eta \in \text{supp}(G_x) : h_x(\eta) < \alpha v\}, \\
D &= \{\eta \in \text{supp}(G_x) : \tilde{h}_x(\eta) < \alpha v\}.
\end{aligned}$$

Suppose $h_x \neq \tilde{h}_x$, then both A and B are not empty by equation (38).

Claim 1: inf A = inf B. Suppose not. Without loss of generality, assume $\inf A > \inf B$.

Then, $\eta^* = \inf A$ satisfies

(a) for any $\eta \in (\underline{\eta}, \eta^*]$, $h_x(\eta) \le \tilde{h}_x(\eta)$;

(b) there exists some $\eta' \in (\underline{\eta}, \eta^*)$ such that $h_x(\eta') < \tilde{h}_x(\eta')$;

(c) $h_x(\eta^*) = \tilde{h}_x(\eta^*) \in (0, \alpha v)$.

From equation (38), we have

$$\underbrace{\int_{\underline{\eta}}^{\eta^*} \left[\frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} - \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)} \right] g_x(\eta) \mathrm{d}\eta}_{\mathrm{LHS}(l)} = \underbrace{\int_{\eta^*}^{\bar{\eta}} \left[\frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)} - \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} \right] g_x(\eta) \mathrm{d}\eta}_{\mathrm{RHS}(l)}$$

However, conditions (a)-(c) imply

$$\lim_{l \to +\infty} \frac{\mathrm{LHS}(l)}{e^{-lh_x(\eta^*)}} = +\infty \tag{39}$$

$$\lim_{l \to +\infty} \frac{\mathrm{RHS}(l)}{e^{-lh_x(\eta^*)}} = 0 \tag{40}$$

where (40) comes from the fact that for any $\eta > \eta^*$, $h_x(\eta) > h_x(\eta^*)$ and $\tilde{h}_x(\eta) > \tilde{h}_x(\eta^*)$, and (39) comes from the fact that, for any $h_x(\eta) < \tilde{h}_x(\eta)$,

$$\frac{\partial}{\partial l} \left[\frac{e^{l(h(\eta^*) - h_x(\eta))}}{\delta + h_x(\eta)} - \frac{e^{l(\tilde{h}(\eta^*) - \tilde{h}(\eta))}}{\delta + \tilde{h}_x(\eta)} \right] > 0$$

$$\frac{\partial^2}{\partial l^2} \left[\frac{e^{l(h(\eta^*) - h_x(\eta))}}{\delta + h_x(\eta)} - \frac{e^{l(\tilde{h}(\eta^*) - \tilde{h}(\eta))}}{\delta + \tilde{h}_x(\eta)} \right] > 0.$$

Equations (39) and (40) contradicts $\mathrm{LHS}(l) = \mathrm{RHS}(l)$ for all $l > 0$.

Claim 2: $\sup C = \sup D$. Let $\eta_1 = \sup C$ and $\eta_2 = \sup_D$. Let

$$\begin{aligned} L_x(l) &:= \int_{\underline{\eta}}^{\bar{\eta}} \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta) \mathrm{d}\eta \\ &= \int_{\underline{\eta}}^{\eta_1} \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta) \mathrm{d}\eta + \frac{e^{-l \cdot \alpha v}}{\delta + \alpha v} (1 - G_x(\eta_1)); \end{aligned}$$

Let $L_x^{(p)}(l)$ denote the pth derivative of $L_x(l)$ with respect to l. Then,

$$\frac{L_x^{(p)}(l)}{(-\alpha v)^p} = \int_{\underline{\eta}}^{\eta_1} \left(\frac{h_x(\eta)}{\alpha v} \right)^p \cdot \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta) \mathrm{d}\eta + \frac{e^{-l \cdot \alpha v}}{\alpha v} (1 - G_x(\eta_1)).$$

Since $0 \le h_x(\eta) < \alpha v$ for all $\eta \in (\underline{\eta}, \eta_1)$, we have

$$\lim_{p\to\infty} \int_{\underline{\eta}}^{\eta_1} \left(\frac{h_x(\eta)}{\alpha v}\right)^p \cdot \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta)\mathrm{d}\eta = 0$$

and

$$1 - G_x(\eta_1) = \alpha v \cdot e^{l\cdot\alpha v} \cdot \lim_{p\to\infty} \frac{L_x^{(p)}(l)}{(-\alpha v)^p}$$

By the same argument, we have

$$1 - G_x(\eta_2) = \alpha v \cdot e^{l\cdot\alpha v} \cdot \lim_{p\to\infty} \frac{L_x^{(p)}(l)}{(-\alpha v)^p}$$

Therefore, $\eta_1 = \eta_2$.

Claim 3: $h_x(\eta^{**}) = \tilde{h}_x(\eta^{**})$ where $\eta^{**} = \sup C = \sup D$. (If $\eta^{**} = \infty$, then $h_x(\eta^{**})$ and $\tilde{h}_x(\eta^{**})$ are understood as $\lim_{\eta\to\infty} h_x(\eta)$ and $\lim_{\eta\to\infty} \tilde{h}_x(\eta)$.)

Suppose, without loss of generality, $h_x(\eta^{**}) > \tilde{h}_x(\eta^{**})$. Then, we must have $\eta^{**} = \bar{\eta}$, since $\eta^{**} < \bar{\eta}$ contradicts to the fact that η^{**} is the supremum of set D. Moreover, we have $\eta^{**} \neq \infty$, otherwise the definition of H requires $\lim_{\eta\to\infty} h_x(\eta) = \lim_{\eta\to\infty} \tilde{h}_x(\eta) = \alpha v$.

Therefore, there exists $\check{\eta} = \sup B$ satisfy

(d) $\check{\eta} < \eta^{**} = \bar{\eta}$;

(e) for all $\eta \in (\check{\eta}, \bar{\eta}]$, $h_x(\eta) > \tilde{h}_x(\eta)$;

(f) $h_x(\check{\eta}) = \tilde{h}_x(\check{\eta}) \in (0, \alpha v)$.

From equation (38), we have, for any $\lambda > 0$

$$\Delta(l) = \int_{\underline{\eta}}^{\check{\eta}} \left[\frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} - \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)}\right] g_x(\eta)\mathrm{d}\eta + \int_{\check{\eta}}^{\bar{\eta}} \left[\frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} - \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)}\right] g_x(\eta)\mathrm{d}\eta = 0$$

Let $\Delta^{(p)}(l)$ denote the pth derivative of $\Delta(l)$. Then, for any $l > 0$

$$\begin{aligned}\frac{\Delta^{(p)}(l)}{(-h_x(\check{\eta}))^p} = & \underbrace{\int_{\underline{\eta}}^{\check{\eta}} \left[\left(\frac{h_x(\eta)}{h_x(\check{\eta})}\right)^p \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} - \left(\frac{\tilde{h}_x(\eta)}{\tilde{h}_x(\check{\eta})}\right)^p \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)}\right] g_x(\eta)\mathrm{d}\eta}_{\mathrm{FI}(l;p)} \\ & + \underbrace{\int_{\check{\eta}}^{\bar{\eta}} \left[\left(\frac{h_x(\eta)}{h_x(\check{\eta})}\right)^p \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} - \left(\frac{\tilde{h}_x(\eta)}{\tilde{h}_x(\check{\eta})}\right)^p \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)}\right] g_x(\eta)\mathrm{d}\eta}_{\mathrm{SI}(l;p)}.\end{aligned}$$

Then, since $h_x(\eta) < h_x(\check{\eta})$ and $\tilde{h}_x(\eta) < \tilde{h}_x(\check{\eta})$ for any $\eta \in (\underline{\eta}, \check{\eta})$, and since $h_x(\eta) > h_x(\check{\eta})$ and $\tilde{h}_x(\eta) > \tilde{h}_x(\check{\eta})$ for any $\eta \in (\check{\eta}, \bar{\eta})$, and since $h_x(\eta) > \tilde{h}_x(\eta)$ for any $\eta \in (\check{\eta}, \bar{\eta})$, we have, for any $l > 0$,

$$\lim_{p\to\infty} \mathrm{FI}(l;p) = 0, \quad \lim_{p\to\infty} \mathrm{SI}(l;p) = +\infty$$

which contradicts to $\Delta^{(p)}(l) = 0$ for all $l > 0$.

Claim 4: $\boldsymbol{h_x \neq \tilde{h}_x}$ leads to contradiction. In other words, we must have $\boldsymbol{h_x = \tilde{h}_x}$.

Let $\eta^* = \inf A = \inf B$ and $\eta^{**} = \sup C = \sup D$. Then, h_x and $\tilde{h}_x$ are strictly increasing and continuously differentiable within $[\eta^*, \eta^{**}]$. Moreover, $h_x(\eta^*) = \tilde{h}_x(\eta^*) = h^*$ and $h_x(\eta^{**}) = \tilde{h}_x(\eta^{**}) = h^{**}$. Therefore,

$$\int_{\eta^*}^{\eta^{**}} \frac{e^{-lh_x(\eta)}}{\delta + h_x(\eta)} g_x(\eta)\mathrm{d}\eta = \int_{h^*}^{h^{**}} e^{-lh}\lambda_x(h)\mathrm{d}h$$
$$\int_{\eta^*}^{\eta^{**}} \frac{e^{-l\tilde{h}_x(\eta)}}{\delta + \tilde{h}_x(\eta)} g_x(\eta)\mathrm{d}\eta = \int_{h^*}^{h^{**}} e^{-lh}\tilde{\lambda}_x(h)\mathrm{d}h$$

where

$$\lambda_x(h) = \frac{g_x(h_x^{-1}(h))}{(\delta + h)\cdot h_x'(h_x^{-1}(h))}$$
$$\tilde{\lambda}_x(h) = \frac{g_x(\tilde{h}_x^{-1}(h))}{(\delta + h)\cdot \tilde{h}_x'(\tilde{h}_x^{-1}(h))}$$

Now, from equation (38), we have

$$\int_{h^*}^{h^{**}} e^{-lh}(\lambda_x(h) - \tilde{\lambda}_x(h))\mathrm{d}h = 0$$

By lemma 7, this equation implies $\lambda_x(h) = \tilde{\lambda}_x(h)$ for any $h \in [h^*, h^{**}]$. Therefore, for any $h \in [h^*, h^{**}]$,

$$G_x(h_x^{-1}(h)) - G_x(\eta^*) = \int_{h^*}^{h} \frac{g_x(h_x^{-1}(h))}{h_x'(h_x^{-1}(h))}\mathrm{d}h = \int_{h^*}^{h} \frac{g_x(\tilde{h}_x^{-1}(h))}{\tilde{h}_x'(\tilde{h}_x^{-1}(h))}\mathrm{d}h = G_x(\tilde{h}_x^{-1}(h)) - G_x(\eta^*)$$

Therefore, $h_x^{-1}(h) = \tilde{h}_x^{-1}(h)$, for any $h \in [h^*, h^{**}]$. Thus, $h_x(\eta) = \tilde{h}_x(\eta)$ for any $\eta \in [\underline{\eta}, \bar{\eta}]$.

By similar arguments, we can prove $h_y = \tilde{h}_y$. Thus, $\{h_x, h_y\}$ is identified. □

Lemma 6. *Let $[a, b]$ be a closed interval where $0 \leq a < b \leq 1$. Suppose $\psi(x)$ is a continuous* *function on* $[a,b]$ *and for all* $n \in \mathbb{N} \cup \{0\}$,

$$\int_a^b x^n \psi(x) \mathrm{d}x = 0.$$

Then, $\psi(x) = 0$ *for all* $x \in [a,b]$.

Proof. Suppose $\psi(x) \neq 0$ for some $x \in [a,b]$. Since $\psi(\cdot)$ is continuous, there exists a closed interval $[c,d]$ where $a < c < d < b$ in which, without loss of generality, $\psi(x)$ is always positive. Let $\beta = \max\{(c-a)(d-a), (b-c)(b-d)\}$, then

$$\forall x \in (c,d), 1 + \frac{1}{\beta}(d-x)(x-c) > 1$$
$$\forall x \in (a,c) \cup (d,b), 0 < 1 + \frac{1}{\beta}(d-x)(x-c) < 1,$$

Thus, the function

$$p(x;r) = \left(1 + \frac{1}{c}(d-x)(x-c)\right)^r$$

has the property that as $r \to \infty$, $p(x;r) \to +\infty$ for $x \in (c,d)$, and $p(x;r) \to 0$ for $x \in [a,c) \cup (d,b]$. Therefore, there exists some $r \in \mathbb{N}$ such that

$$\int_a^b p(x;r)\psi(x) \mathrm{d}x > 0$$

On the other hand, $p(x;r)$ is a polynomial of order $2r$. Since for $x \in \mathbb{N} \cup \{0\}$

$$\int_a^b x^n \psi(x) \mathrm{d}x = 0,$$

we should have, for any $r \in \mathbb{N}$,

$$\int_a^b p(x;r)\psi(x) \mathrm{d}x = 0$$

Contradiction! Therefore, $\psi(x) = 0$ for $x \in [a,b]$. □

Lemma 7. *Let* $[a,b]$ *be a closed interval where* $0 \leq a < b$ *, and* $\lambda(x)$ *be a continuous function on* $[a,b]$*. Suppose for any* $l > 0$,

$$\int_a^b e^{-lx} \lambda(x) \mathrm{d}x = 0.$$

Then, $\lambda(x) = 0$ *for any* $x \in [a, b]$*.*

Proof. Let $y = e^{-x}$. Then,

$$\int_a^b e^{-lx}\lambda(x)\mathrm{d}x = 0 \Leftrightarrow \int_{e^{-b}}^{e^{-a}} y^{l-1}\lambda(-\log(y))\mathrm{d}y = 0$$

Therefore, $\lambda(-\log(y)) = 0$ for all $y \in [\exp(-b), \exp(-a)]$ by lemma 6. Thus, $\lambda(x) = 0$ for all $x \in [a, b]$. $\square$

References

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Disentangling the Preferences of Teachers and Employers," *Journal of Labor Economics*, January 2013, *31* (1), 83–117.

Bruze, Gustaf, Michael Svarer, and Yoram Weiss, "The Dynamics of Marriage and Divorce," *Journal of Labor Economics*, January 2015, *33* (1), 123–170.

Chiappori, Pierre-André and Bernard Salanié, "The econometrics of matching models," *forthcoming Journal of Economic Literature*, 2015.

_ **, Sonia Oreffice, and Climent Quintana-Domeque**, "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market," *Journal of Political Economy*, August 2012, *120* (4), 659–695.

Choo, Eugene, "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 2015, *83* (4), 1373–1423.

_ **and Aloysius Siow**, "Who marries whom and why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

_ **and** _ , "Lifecycle marriage matching: Theory and Evidence," in "2007 Meeting Papers" Society for Economic Dynamics 2007.

Eeckhout, Jan and Philipp Kircher, "Identifying Sorting—In Theory," *The Review of Economic Studies*, July 2011, *78* (3), 872–906.

Flinn, Christopher and James Heckman, "New methods for analyzing structural models of labor force dynamics," *Journal of Econometrics*, January 1982, *18* (1), 115–168.

Fox, Jeremy, "Identification in matching games," *Quantitative Economics*, November 2010, *1* (2), 203–254.

_ **and Patrick Bajari**, "Measuring the Efficiency of an FCC Spectrum Auction," *American Economic Journal. Microeconomics*, February 2013, *5* (1), 100–146.

_ **, David Hsu, and Chengyu Yang**, "Unobserved Heterogeneity in Matching Games with an Application to Venture Capital Investments," *Working Paper*, 2015.

Galichon, Alfred and Bernard Salanié, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *Working Paper*, 2014.

Goussé, Marion, "Marriage Gains: Who Should You Marry ?," SSRN Scholarly Paper ID 2533736, Social Science Research Network, Rochester, NY May 2014.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii, "Identifying Equilibrium Models of Labor Market Sorting," Working Paper 896, Society for Economic Dynamics 2014.

Hanche-Olsen, Harald and Helge Holden, "The Kolmogorov–Riesz compactness theorem," *Expositiones Mathematicae*, 2010, *28* (4), 385–394.

Jacquemet, Nicolas and Jean-Marc Robin, "Assortative matching and search with labour supply and home production," *Working Paper*, 2012.

Lamadon, Thibaut, Jeremy Lise, Costas Meghir, and Jean-Marc Robin, "Matching, Sorting, and Wages," *Working Paper*, 2013.

Rudin, Walter, *Functional analysis* International series in pure and applied mathematics, 2nd ed ed., New York: McGraw-Hill, 1991.

Shimer, Robert and Lones Smith, "Assortative matching and search," *Econometrica*, 2000, *68* (2), 343–369.

Shin, Seul Ki, "Preferences vs. Opportunities: Racial/Ethnic Intermarriage in the United States," SSRN Scholarly Paper ID 2526169, Social Science Research Network, Rochester, NY November 2014.

Stokey, Nancy L., Robert E. Lucas, and Edward C. Prescott, *Recursive methods in economic dynamics*, Cambridge, Mass: Harvard University Press, 1989.

Tröger, Thomas and Georg Nöldeke, "Matching Heterogeneous Agents with a Linear Search Technology," *Working Paper*, 2009.

Wong, Linda Y, "Structural Estimation of Marriage Models," *Journal of Labor Economics*, 2003, *21* (3), 699–727.