WORKING PAPERS

N° 18-897

March 2018

"Identification in One-to-One Matching Models with Nonparametric Unobservables"

Shruti Sinha

Toulouse School of Economics

Identification in One-to-One Matching Models with Nonparametric Unobservables*

Shruti Sinha†

March, 2018

Abstract

This paper considers a one-to-one matching model with transferable utilities, in two-sided markets. In the model, the agents have preferences over some observable agent characteristics (called types) on the other side of the market. There are other observed characteristics aggregated at the level of types that determine the systematic preferences over these types. These systematic preferences enter the agent utilities in the form of a linear index. Agents also have idiosyncratic taste shocks. This paper shows the identification of systematic preference parameters over types, without making any parametric assumptions on the distribution of the unobserved taste shocks. The matching model reduces to two separate discrete-choice problems linked together by market clearing conditions, satisfied in the presence of equilibrium transfers. However, transfers are endogenous and unobserved which makes the discrete-choice problem non-standard. This paper gives conditions under which transfers are simply functions of the linear indices. This insight along with variation across i.i.d. markets is used to reduce the matching model to a semiparametric multi-index model with an unknown link function. Identification is shown under appropriate exclusion restrictions on the regressors.

Keywords: One-to-One Matching, Transfers, Identification

JEL: C31, C35, C78, J12

*I am deeply indebted to my advisors Elie Tamer and Ivan Canay, for their extensive guidance and encouragement. I am also grateful to my committee member, Alex Torgovitsky, who has been generous with his advice and feedback. This paper has also benefited from discussions with Christian Bontemps, Yeon-Koo Che, Alfred Galichon, Eric Gautier, Yinghua He, Joel Horowitz, Thierry Magnac, Chris Rose, Bernard Salanié, Arkadius Szydlowski, Scott Thompson, and seminar participants at Northwestern University, University of Illinois - Urbana Champaign, Toulouse School of Economics, University of Leicester, French Econometrics Conference 2015, European Winter Meetings 2015, SETA/NZESG 2016, Royal Economic Society Conference 2016, Bristol Econometrics Study Group 2016, University of Melbourne-TSE joint conference 2016.

†Toulouse School of Economics, Université Toulouse 1 - Capitole. Email: shruti.sinha@tse-fr.edu.

1 Introduction

Matching markets are two-sided markets, where agents on both sides of the market have preferences over forming matches with agents on the other side. Empirical studies of matching markets have recently garnered considerable attention. Following the pioneering contribution of Becker (1973), many such studies view matching outcomes as a competitive equilibrium with transferable utilities, and assume that the matching process is based on exactly one characteristic. Becker's one-dimensional model implies that agents with similar characteristics will find it optimal to match. Typically, a researcher observes a pattern of how the agents sort themselves into matches. These patterns can be based on multiple characteristics and we do not always observe agents with similar characteristics matching. Thus, the one-dimensional Beckerian model of matching is too simplistic to explain the different match patterns that are actually observed in the data. An alternative is to allow preferences to depend on multiple characteristics, including those that are latent (or private to the agent). The research goal is to then identify the underlying preferences that explain observed match patterns. Estimates of these preferences can be used to understand the relative importance of different characteristics in the equilibrium matching of agents (Chiappori, Oreffice, and Quintana-Domeque 2012), to estimate the gains to matching (Botticini and Siow 2008), or to do counterfactual analysis (Chiappori, Salanié, and Weiss 2017).

This paper contributes to a growing literature on the identification and estimation of preference parameters in one-to-one matching models with transferable utilities. Typically, matching markets have two disjoint groups of agents (called "sides" of the market). Agents on each side of the market can either form a match with exactly one agent on the other side, or remain unmatched. On each side, agents are divided into a finite number of groups called "types". A type corresponds to one or more characteristics of agents over which we observe sorting. Agents on each side have preferences defined over the set of types, instead of over the set of all agents, on the other side. A match between two agents, say agent k of type i on one side and agent l of type j on the other side, is formed when both agents find forming the match utility-maximizing. We allow preferences to include transferable utilities. Transfers act as prices which are determined simultaneously with the match allocation, and adjust in a way that each agent maximizes their utility and the market clears. In particular, we impose the following structure on utilities,

$$\begin{aligned} \text{Side } \mathbf{a}: \ u_{k,ij} &= X_{ij}\alpha_i + \varepsilon_{kj} - \tau_{ij}, \\ \text{Side } \mathbf{b}: \ v_{l,ij} &= Z_{ij}\beta_j + \nu_{il} + \tau_{ij}, \end{aligned} \tag{1}$$

where $u_{k,ij}$ is utility of agent k of type i when matched with an agent of type j on the other side, and $v_{l,ij}$ is utility of agent l of type j when matched with agent of type i on the other side. Here, $X_{ij}\alpha_i$, $Z_{ij}\beta_j$ are linear indices which capture the systematic part of the preferences of agents matching; $\varepsilon_{kj}, \nu_{il}$ are idiosyncratic taste shocks specific to each agent; and τ_{ij} are equilibrium transfers associated with each type of match. The match allocation supported by these equilibrium transfers will be unique, efficient, and pairwise stable. These properties of one-to-one matching with transfers have been studied in Koopmans and Beckmann (1957), Shapley and Shubik (1972), Becker (1973, 1974), Gretsky, Ostroy and Zame (1992, 1999) and have been summarized in Roth and Sotomayor (1990, Chapter 8).

The preference structure in (1) is similar to the ones studied in Choo and Siow (2006) and Galichon and Salanié (2015). As in their models, agents' preferences are composed of three additively separable parts that depend solely on the type of match being formed, transfers, and an idiosyncratic taste shock. We deviate from the Choo-Siow and Galichon-Salanié framework in the following three ways. First, we explicitly model the systematic part of the preferences as a function of some covariates that characterize the types. Second, we make no parametric assumptions on the distribution of the taste shocks, ε, ν. Third, we exploit variation across multiple similar markets for identification.

Our goal is to identify and estimate the systematic part of preferences, which in our model corresponds to the coefficients, α_i, β_j, of the linear index that enters agent utilities. Since, in equilibrium, every agent is maximizing preferences over a finite number of alternatives, we can view the matching process as a discrete-choice problem on the two sides. There is a large literature on the identification and estimation of such parameters in discrete-choice models. However, the discrete-choice problems in our framework are non-standard as they depend on transfers, τ_{ij}, which are not observed by the researcher. In this paper we view transfers as functions of type attributes (X_{ij} and Z_{ij}), where a transfer function maps observed covariates to equilibrium transfers. We show that the transfer functions depend on these characteristics only through the linear indices. With this insight we are able to reduce the matching model in many markets to a conditional mean model which resembles a more familiar problem in econometrics.

In each market, the discrete-choice problem generates a vector of match proportions (for every type of match that can be formed) as functions of the linear index, transfer function, and distribution of unobserved taste shocks. Aggregating this information across all markets and using our result on transfer functions, we are able to reduce this to a multiple index model with unknown link functions and multiple equations. This approach is new in matching models.

We show that the coefficient of the indices is identified up to a scaled normalization under some exclusion restrictions, and smoothness and non-linearity restrictions on the unknown link function. These assumptions are similar to the ones made in Ichimura and Lee (1991), which studies identification in a semiparametric multi-index model with a single equation. Here we modify the assumptions to accommodate for multiple equations.

The literature on identification of preferences in matching models with transferable utility can be broadly divided into two strands based on whether the choice sets of agents consist of types or individual agents. The one-to-one matching model with transferable utility has mostly been studied where agents choose types rather than individual agents, where agents of the same type are assumed to be exchangeable. Among these are Echenique et al. (2013), Dupuy and Galichon (2014), Galichon and Salanié (2015), Graham (2013b) among others, which build on the seminal work of Choo and Siow (2006). Choo and Siow (2006; henceforth, the CS model) consider identification of the systematic preference parameters as solely a function of the match pattern when the number of types is finite. However, this approach does not explain why such preferences occur. In this paper, we model what factors determine these preferences by introducing covariates in the systematic part of the utility.

The CS model also allows for matching to occur on unobserved idiosyncratic taste shocks that can be different for every agent. They identify systematic preferences by assuming unobserved taste shocks to be distributed as extreme value type I. This model is extended in Graham (2013b), where the unobservables are still assumed to have an extreme value distribution but an unknown scaling parameter is added to this distribution on both sides. Galichon and Salanié (2015) also study the CS model where the unobservables can belong to a richer class of distributions, but where the distribution still needs to be known *a priori.* Echenique et al. (2013) also consider identification in matching models with transferable utility, but their model differs from the CS model as it allows for only aggregate taste shocks (identical shocks for all individuals with the same type). Galichon and Dupuy (2014) extends the CS model to a continuous logit model. All of these papers use a single large market for identification. In this paper, we make no parametric assumptions on the distribution of the unobserved taste shock. Thus, in contrast to the existing literature, we cannot rely on the shape of these distributions to yield identification. Instead, we use variation across many markets in the covariates, introduced in the systematic preferences, for identification.

On the other hand, Fox (2010), Chiappori, Oreffice, and Quintana-Domeque (2012), and Fox, Hsu, and Yang (2017) consider identification of features of the match surplus function when the choice set consists of individual agents. Agents are not exchangeable in this context. Fox (2010) relies on a "rank order" assumption, and Fox, Hsu, and Yang (2017) allow for match preferences to be based on observable and unobservable characteristics. However, their objective is complementary to ours, in that they are interested in identifying features of the distribution of the unobservables. Finally, Chiappori, Oreffice, and Quintana-Domeque (2012) present a general model of matching with unobserved heterogeneity, but recover the systematic part of the preferences only ordinally. All these papers observe matching data from multiple markets.

The CS framework has been used in many empirical papers that have studied marriage patterns, see Fox (2009) for a review. For example, Choo and Siow (2006) used it to link changes in gains to marriage and abortion laws. Siow and Choo (2006) apply this model to measure the impact of demographic changes on matching patterns. Botticini and Siow (2008) and Chiappori, Salanie, and

Weiss (2017) use this model to study returns to education on the marriage market. Banerjee et al. (2013) estimates the impact of caste and non-caste attributes in determining matching patterns in India. Although marriage markets are a leading application of this setting, there are other markets that can be modeled in this framework, for example, the matching of firms with CEOs (Tervio 2008, and Gabaix and Landier 2008).

There is also a recent related literature that studies identification of preferences with match outcomes in a one-to-one non-transferable utilities framework (see for example Dagsvik 2000, Agarwal and Diamond 2017, and Menzel 2015). In these models, the choice sets of agents are endogenous, making the discrete-choice approach non-standard. In this setting, pairwise stable match allocations are not unique (in contrast to transferable utility models, where match allocations are unique). Finally, Galichon, Kominers, and Weber (2016) consider one-to-one matching in imperfectly transferable utility models which differ from TU models in that a pairwise stable outcome may no longer coincide with the efficient outcome. See Chiappori and Salanié (2016) for a recent survey of matching models in econometrics.[1]

The remainder of the paper is organized as follows. Section 2 sets up the matching framework and the notation. We explore the role of transfers and present the main result on transfers in Section 2.2. We use this result in Section 3 to aggregate matching information over many markets and reduce the problem to a semiparametric multi-index model. Section 4 discusses the assumptions under which the parameters (coefficients of the linear indices) in the reduced form model are identified. Finally, Section 5 concludes. All proofs are in the Appendix.

2 The Model

Let $\mathbf{T} = \{1, \cdots, T\}$ be the set of all markets. Each market $t \in \mathbf{T}$ is a matching market composed of two sides, $\mathbf{a}$ and $\mathbf{b}$. For any market t, let $\mathbf{N}_{\mathbf{a}t}$ be the set of all agents on side $\mathbf{a}$ and $\mathbf{N}_{\mathbf{b}t}$ be the set of all agents on side $\mathbf{b}$. We will not allow any overlap between these two sets of agents; that is, $\mathbf{N}_{\mathbf{a}t} \cap \mathbf{N}_{\mathbf{b}t} = \varnothing$. Agents on each side will be characterized by their *types*. A type denotes some characteristic(s) that are common to a subset of agents on a given side of the market. Let $\mathbf{I} = \{1, \cdots, I\}$ be the finite set of types on side $\mathbf{a}$ and $\mathbf{J} = \{1, \cdots, J\}$ be the finite set of types on side $\mathbf{b}$.

To assign types to agents we introduce the functions

$$\mathbf{i} : \mathbf{N}_{\mathbf{a}t} \to \mathbf{I},$$

$$\mathbf{j} : \mathbf{N}_{\mathbf{b}t} \to \mathbf{J}.$$

[1] Graham (2011, 2013a) provide older surveys on this topic.

Note $\mathbf{I}$ and $\mathbf{J}$ stay the same across every market. In any market $t \in \mathbf{T}$, let $s_{\mathbf{a}it}$ denote the proportion of type i agents on side $\mathbf{a}$, and $s_{\mathbf{b}jt}$ denote the proportion of type j agents on side $\mathbf{b}$. The vector $S_t = (s_{\mathbf{a}1t}, \cdots, s_{\mathbf{a}It}, s_{\mathbf{b}1t}, \cdots, s_{\mathbf{b}Jt})$ then characterizes the proportions of types in market t.

In a one-to-one matching model each agent on side $\mathbf{a}$ is allowed to match with exactly one agent on side $\mathbf{b}$, or agents can remain unmatched. If agent $k \in \mathbf{N_{at}}$ matches with agent $l \in \mathbf{N_{b}}_t$, then the match pair will be denoted by (k, l). All match pairs (k, l) such that $\mathbf{i}(k) = i$ and $\mathbf{j}(l) = j$ are said to be of *match type* (i, j). If $k \in \mathbf{N_{at}}$ where $\mathbf{i}(k) = i$ remains unmatched, we will denote the match pair by $(k, 0)$ and the match type by $(i, 0)$. Similarly, if $l \in \mathbf{N_{bt}}$ where $\mathbf{j}(l) = j$ remains unmatched, we will denote the match pair by $(0, l)$ and the match type by $(0, j)$. Let $\mathbf{M}$ be the collection of all match types where,

$$\mathbf{M} = \{(i, j) : i \in \mathbf{I} \cup \{0\}, j \in \mathbf{J} \cup \{0\}, (i, j) \neq (0, 0)\}.$$

Any agent has preferences over the set of types on the other side of the market. These preferences are determined by a systematic part that is common to agents of the same type, and an idiosyncratic part that can be different for each agent. For instance, an agent $k \in \mathbf{N_{at}}$ with $\mathbf{i}(k) = i$ has preferences over the set $\mathbf{J} \cup \{0\}$ composed of a systematic part that only depends on its type i and a vector of the agent's idiosyncratic taste shocks, ε_{kt}. We model the systematic part of the utility for a match as functions of covariates (denoted by X_{ijt}) that characterize the types. That is, X_{ijt} captures the various attributes of a match type (i, j) that agents of type i care about. These covariates have to vary with types and market. The vector of idiosyncratic taste shocks for agent k of type i is given by $\varepsilon_{kt} = (\varepsilon_{k0t}, \varepsilon_{k1t}, \varepsilon_{k2t}, \cdots, \varepsilon_{kJt})$. This is drawn from a distribution F_i for each $i \in \mathbf{I}$, where ε_{kjt} denotes k's personal taste to match with type j agents on the other side, $j = 1, \cdots, J$, and ε_{k0t} denotes k's personal taste to remain unmatched. Similarly, an agent $l \in \mathbf{N_{bt}}$ with $\mathbf{j}(l) = j$ will have preferences over the set $\mathbf{I} \cup \{0\}$ based on a systematic part that is a function of covariates, Z_{ijt}, and the agent's idiosyncratic taste shock ν_{lt}, where $\nu_{lt} = (\nu_{0lt}, \nu_{1lt}, \cdots, \nu_{Ilt})$. This is drawn from a distribution G_j for all $j \in \mathbf{J}$. We will assume that, for all i and j, F_i and G_j are the same across all markets.

Let the number of (i, j) matches in market t be denoted by π_{ijt}, for any $(i, j) \in \mathbf{M}$. Then define the match allocation in market t to be the following matrix,

$$\Pi_t = \begin{pmatrix} \pi_{11t} & \pi_{12t} & \cdots & \pi_{1J} & \pi_{10t} \\ \pi_{21t} & \pi_{22t} & \cdots & \pi_{2Jt} & \pi_{20t} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \pi_{I1t} & \pi_{I2t} & \cdots & \pi_{IJt} & \pi_{I0t} \\ \pi_{01t} & \pi_{02t} & \cdots & \pi_{0Jt} & - \end{pmatrix},$$

where the $(I + 1, J + 1)th$ element has been removed as it has no meaning.

The match allocation in each market t will be viewed as the *competitive equilibrium* of an economy with transferable utilities.[2] A competitive equilibrium will be a match allocation and an $I \times J$ matrix of *equilibrium transfers* Γ_t, where

$$\Gamma_t = \begin{pmatrix} \tau_{11t} & \cdots & \tau_{1Jt} \\ \vdots & \ddots & \\ \tau_{I1t} & & \tau_{IJt} \end{pmatrix}.$$

The transfers τ_{ijt} will only depend on the match type (i,j) and not on the identity (k,l) of the match (this will follow from arguments in Salanié (2015)).

In a competitive equilibrium each match pair will share a *match surplus* according to the equilibrium transfers. The share of agents k and l in an (i,j) match type are given by

$$\begin{aligned} U_{k,ijt} &= X_{ijt}\alpha_i + \varepsilon_{kjt} - \tau_{ijt}, \\ U_{l,ijt} &= Z_{ijt}\beta_j + \nu_{ilt} + \tau_{ijt}, \end{aligned} \tag{2}$$

where $\mathbf{i}(k) = i$ and $\mathbf{j}(l) = j$. Here X_{ijt} and Z_{ijt} represent d_x and d_z dimensional vectors of covariates (that characterize match types), respectively, for every match type $(i,j) \in \mathbf{I} \times \mathbf{J}$ and every market $t \in \mathbf{T}$. These covariates enter through linear indices with coefficients $\alpha_i \in \mathbb{R}^{d_x}$ for every $i \in \mathbf{I}$ and $\beta_j \in \mathbb{R}^{d_z}$ for every $j \in \mathbf{J}$. The linear indices represent the systematic part of the preferences. As introduced above, ε_{kjt} denotes the taste shock for agent k on side $\mathbf{a}$ for type j agents on side $\mathbf{b}$; ν_{ilt} denotes the taste shock for agent l on side $\mathbf{b}$ for type i agents on side $\mathbf{a}$. If agents remain unmatched, there are no transfers to be exchanged, and we normalize the linear index component to be zero. Then utility of unmatched agents $k \in \mathbf{N}_{\mathbf{a}t}$ of match type $(i,0)$ and $l \in \mathbf{N}_{\mathbf{b}t}$ of match type $(0,j)$ are given by

$$U_{k,i0t} = \varepsilon_{k0t} \quad \text{and} \quad U_{l,0jt} = \nu_{0lt},$$

respectively.

In any market $t \in \mathbf{T}$, a ***competitive equilibrium*** will be defined as a set of transfers $\mathbf{\Gamma}_t$ and a match allocation Π_t such that every agent maximizes their utility and the market clears. The complete choice set of agents on side $\mathbf{a}$ is given by $\mathbf{J} \cup \{0\}$ and for agents on side $\mathbf{b}$ is given by

[2]The usual notion of equilibrium in matching games is that of pairwise stability. In one-to-one transferable utility models the match allocation in a pairwise stable match outcome coincides with the match allocation in a competitive equilibrium (Shapley and Shubik 1972). Moreover, when markets are large (an assumption that we formalize in Section 2.3) there are unique transfers (or, prices) that support the equilibrium match allocation (Gretsky, Ostroy, and Zame (1992,1999)).

$\mathbf{I} \cup \{0\}$. Then in equilibrium, for a match (k, l) of match type $(i, j) \in \mathbf{M}$ to occur it must be that

$$j \in \arg \max_{j' \in \mathbf{J} \cup \{0\}} U_{k,ij't},$$
$$i \in \arg \max_{i' \in \mathbf{I} \cup \{0\}} U_{l,i'jt}. \quad (3)$$

and the markets clear,

$$\sum_{j=0}^{J} \pi_{ijt} = s_{\mathbf{a}it} |\mathbf{N}_{\mathbf{a}t}|, \quad \forall\, i \in \mathbf{I},$$
$$\sum_{i=0}^{I} \pi_{ijt} = s_{\mathbf{b}jt} |\mathbf{N}_{\mathbf{b}t}|, \quad \forall\, j \in \mathbf{J}. \quad (4)$$

The equilibrium match allocation will be unique even if transfers are not (see Shapley and Shubik 1972; Gretsky, Ostroy, and Zame 1992).

2.1 Example: A Marriage Market

In this subsection we will describe a marriage market in the setting of our model. In the rest of the paper we will use this example to illustrate the main assumptions and the intuition for the results.

Consider a marriage market where the set of men and women form two sides of a market. In this context we can think of a market to be a city. To keep with the application in Choo and Siow (2006), let types be given by age on both sides. Choo and Siow let types be given by the different ages between 16 and 75. Some of their findings include that a 20-year-old man prefers to match with a slightly younger woman, and a 20-year-old woman prefers to match with a slightly older man. On the other hand, 40-year-old men and women have more dispersed preferences.[3] This paper suggests an approach to further understanding these preferences.

Here, for the ease of exposition, suppose there are only two age types on each side

$$\mathbf{I} = \{20, 40\}, \quad \mathbf{J} = \{20, 40\}.$$

The different match types in set $\mathbf{M}$ will be given by the following cells:

[3] Choo and Siow (2006) measure preferences as distribution of systematic net gains from marriage. These are quantities of the type $\log(\pi_{ij}/\pi_{i0})$ which denotes net gains of a type i man from matching a type j woman relative to remaining unmatched. Similarly, $\log(\pi_{ij}/\pi_{0j})$ denotes net gains of a type j woman from matching a type i man relative to remaining unmatched.

		Women		
		Age 20	Age 40	
	Age 20	$(20, 20)$	$(20, 40)$	$(20, 0)$
Men	Age 40	$(40, 20)$	$(40, 40)$	$(40, 0)$
		$(0, 20)$	$(0, 40)$	-

Suppose agents' preferences over types are based on expected education levels and the expected income of their spouse. Moreover, agents care about having similar expected education levels as their spouse.[4] Then covariates X_{ijt} can contain variables $(X_{ijt,educ}, X_{jt,inc})$ where $X_{ijt,educ} = |X_{it,educ} - X_{jt,educ}|$ denotes the difference in expected education between a type i (i.e. age i) man and a type j (i.e. age j) woman, and $X_{jt,inc}$ denotes the expected income for a type j woman. Similarly, we can define covariates $Z_{ijt} = (Z_{ijt,educ}, Z_{it,inc})$. The characteristics in X_{ijt} captures the attributes of match type (i, j) and the type of match partner j, and Z_{ijt} captures the attributes of match type (i, j) and type of match partner i.

The systematic part of the preferences of a man of type i for a woman of type j is given by index $X_{ijt}\alpha_i$ where $\alpha_i = (\alpha_{i,educ}, \alpha_{i,inc})'$. Similarly, the systematic part of the preferences of a woman of type j for a man of type i is given by index $Z_{ijt}\beta_j$ where $\beta_j = (\beta_{j,educ}, \beta_{j,inc})'$. These coefficients act as weights that the agents put on the different type characteristics.

We also introduced the variable S_t which represented the proportions of types. The role of this is explained in Section 2.3. Suppose that the population of men in some market t (which is a city) is composed 60% of 20-year-olds and 40% of 40-year-olds, and the population of women is composed of 50% of 20-year-olds and 40-year-olds. Then, $S_t = (0.6, 0.4, 0.5, 0.5)$.

An important distinction from the CS framework is that, here, types i, j and covariates X_{ijt} and Z_{ijt} are different objects. In the CS model, including additional covariates corresponds to increasing the number of types. For example, consider including agent education in the marriage example above which takes only two values $\{High, Low\}$, say. Then the number of types of the form (age, education) increases from two to four on each side,

$$\mathbf{I} = \{(20, H), (20, L), (40, H), (40, L)\}, \ \mathbf{J} = \{(20, H), (20, L), (40, H), (40, L)\}.$$

Also all possible match types increase from 8 cells to 24 cells. So making agent preferences depend on several factors like income, education, age, etc. will increase the number of match types and typically leads to the practical problem of having many empty (or thin) cells in the data.[5] In contrast, in

[4]In practice, we can measure expected income (education) of an agent of type i as the average income (education) of type i agents. Instead of expectations (or, averages) we could have considered any aggregate measure of the variables education and income that could characterize a match type (i, j).

[5]This is a common problem in empirical discrete-choice problems and also encountered in the empirical section of Choo and Siow (2006; see footnote 15).

the setup described in this paper, we have to first take a stand on which characteristic(s) will define types (e.g. age), and then use these types to aggregate other characteristics (e.g. education, income) that agents may care about. This has two implications different from the CS model. First, we can make the systematic part of the preferences depend on a rich set of covariates without increasing the number of elements in the set of match types, $\mathbf{M}$. In this sense, we are carrying out a dimension reduction of the set $\mathbf{M}$ in the CS framework. Second, it imposes structure on the systematic part of preferences over types, as functions of other characteristics. Thus, we can study what determines preferences over types.

2.2 Discussion on Agent Preferences

As in the CS framework, agents on one side of the market have preferences only over the type of agents on the other side and not on the individual identity of the agent. That is, any $k \in \mathbf{N_{a}}_t$ on side $\mathbf{a}$ ranks the elements of set $\mathbf{J} \cup \{0\}$ and not $\mathbf{N_{b}}_t$. Similarly, any $l \in \mathbf{N_{b}}_t$ on side $\mathbf{b}$ ranks the elements of set $\mathbf{I} \cup \{0\}$. This is captured by the utility specification in (2) where the utility of agent k from match (k, l) of type (i, j) does not depend on l but only on type j. Similarly, the utility for l only depends on type of match partner i and not on k. An implication of this is that agents of the same type act as *perfect substitutes* (or, are exchangeable) for agents on the other side. For instance, suppose $k \in \mathbf{N_{a}}_t$ has most preferred type j on side $\mathbf{b}$, and let $l, l' \in \mathbf{N_{b}}_t$ be such that $\mathbf{j}(l) = \mathbf{j}(l') = j$. Then, k is indifferent between matching with l and l'. Thus, even though the model cannot capture why agents on side $\mathbf{a}$ match with agent l versus agent l', it does capture why they choose type j versus type j'. In the marriage example in Section 2.1, it means that each man (woman) treats women (men) of a given type as exchangeable.

The systematic part of preferences over types are determined by covariates X_{ijt} and Z_{ijt} which are aggregated at the level of types. In the marriage example, one of the group attributes for, say, 20-year-old women is $X_{20t,inc}$ which denotes the *expected* income of 20-year-old women in the market. Also suppose $X_{20,20,t}\alpha_{20} > X_{20,40,t}\alpha_{20} > X_{20,0,t}\alpha_{20}$, then the systematic preference of any 20 year-old-man is given by: *marry a 20-year-old woman* $\succ$ *marry a 40-year-old woman* $\succ$ *remain unmatched*.

Further, agents of the same type can have different preferences (*heterogenous tastes*) due to the presence of idiosyncratic taste shocks, ε_{kj} and ν_{il}. This is even though the agents of a given type have the same systematic preferences. For example, consider two 20-year-old men k and k'. Then ε_{kt} and $\varepsilon_{k't}$ can take realizations such that $X_{20,20,t}\alpha_{20} + \varepsilon_{k,20,t} > X_{20,40,t}\alpha_{20} + \varepsilon_{k,40,t}$ and $X_{20,20,t}\alpha_{20} + \varepsilon_{k',20,t} < X_{20,40,t}\alpha_{20} + \varepsilon_{k',40,t}$. Then, ignoring transfers, k will prefer to marry a 20-year-old woman, whereas k' will prefer to marry a 40-year-old woman. We aggregate over these idiosyncratic taste shocks and use exchangeability of agents of the same type to compute market shares of matches in each market, in Section 3.

It is worth observing that we allow α_i to differ with i and β_j to differ with j. This allows for heterogeneity across types in the systematic part of preferences and does not rule out match preferences where, for example, older (younger) types systematically prefer to match with older (younger) types. This flexibility seems important to capture realistic patterns of preferences.

Of course, an important choice in our framework is which attributes of an individual should determine the types, and which should be captured in the covariates, X_{ijt}, Z_{ijt}. In practice, a range of factors may determine this choice. One is the dimension reduction aspect discussed above. Another may reflect the researcher's understanding of the matching process. For instance, types may be attributes observable to agents at the time of the match (e.g. ethnicity, age) while the covariates may be unobserved but inferred based on the observable types (e.g. income).

2.3 Transfers

Transfers play an important role in our model, as they ensure markets clear.[6] With transfers entering the utility specification directly, we can treat all agents as maximizing their preferences over the entire type set and their outside option to remain unmatched. This is simply a discrete-choice model on both sides in each market. The complication here is that transfers are endogenous and not observed. Thus, we are unable to use results directly from the discrete-choice literature.

In the existing matching literature with transferable utilities, there are two approaches to handling transfers. The first approach relies on functional form assumptions on the distribution of ε and ν. For instance, if we assume (as in Choo and Siow, 2006) that, for any agent $k \in \mathbf{N_a}_t$ and $l \in \mathbf{N_b}_t$, $\{\varepsilon_{kj} : j \in \mathbf{J} \cup \{0\}\}$ and $\{\nu_{il} : i \in \mathbf{I} \cup \{0\}\}$ are drawn independently from the extreme value type I distribution, then the systematic part of the match surplus is identified as $X_{ij}\alpha_i + Z_{ij}\beta_j = \log(\pi_{ij}/\pi_{i0}) + \log(\pi_{ij}/\pi_{0j})$, for all (i, j). As Choo and Siow (2006) show, this follows by eliminating transfers using the logit specification. Here, we instead allow the joint distribution of $(\varepsilon_{k0}, \varepsilon_{k1}, \cdots, \varepsilon_{kJ}) \sim F_i, \mathbf{i}(k) = i$, to be unknown and have an arbitrary correlation structure. We also allow F_i to be different for different types $i \in \mathbf{I}$. The same is true for each G_j. Therefore, here we cannot use the shape of the distribution functions to eliminate transfers.

The second approach to deal with transfers exploits the efficiency property of the equilibrium, which says that the match allocation Π_t maximizes the total match surplus in the market t. Thus, one only needs to consider the sum of total match surplus in the market, which does not involve any transfer terms. Fox (2010) assumes a rank-order property to nonparametrically identify the match surplus function. Galichon and Salanié (2015) consider the surplus maximization problem as a convex optimization problem. This determines the match surplus in terms of distributions F_i and G_j (for all i and j) which are assumed to be known a priori. Their approach does not apply to settings where F_i and G_j are unknown.

[6] For a discussion on having transferable utilities for marriage see Becker (1973).

In this section we will show that, under some assumptions, transfers are functions of the linear indices introduced in equations (2). This gives us an alternative way to deal with transfers which is new in the literature.

Let

$$X_t = (X_{11t}, X_{12t}, \cdots, X_{1Jt}, X_{21t}, X_{22t}, \cdots, X_{IJt}) \in \mathcal{X}^{IJ},\ \mathcal{X} \subseteq \mathbb{R}^{d_x},$$

$$Z_t = (Z_{11t}, Z_{21t}, \cdots, Z_{I1t}, Z_{12t}, Z_{22t}, \cdots, Z_{IJt}) \in \mathcal{Z}^{IJ}, \quad \mathcal{Z} \subseteq \mathbb{R}^{d_z},$$

be the lists of all covariate vectors in a market t. A market's characteristics are then summarized by (Π_t, X_t, Z_t, S_t), where recall that Π_t is the equilibrium match allocation and S_t gives the proportions of types. We impose the following assumptions on the sample of markets from which the data is drawn.

First, we make a large market assumption common in the literature which will guarantee the uniqueness of equilibrium transfers in each market. This will involve assuming that sets $\mathbf{N}_{\mathbf{a}t}$ and $\mathbf{N}_{\mathbf{b}t}$ for each market t are masses of infinitesimal agents (see Galichon and Salanié 2015, for a similar assumption).

Assumption 2.1 *In each market t, agents are infinitesimal and the total mass of the population is normalized to 1, with $\sum_{i=1}^{I} s_{\mathbf{a}it} + \sum_{j=1}^{J} s_{\mathbf{b}jt} = 1$.*

Our approach is intended to exploit data on many markets by making use of the following further assumption.

Assumption 2.2 *(a) The distributions of shocks, as given by $\{F_i : i \in \mathbf{I}\}$ and $\{G_j : j \in \mathbf{J}\}$ for all i and j, do not depend on the market t nor on the realization of market covariates X_t and Z_t.*

(b) For all i and j, F_i and G_j are absolutely continuous and have full support on $\mathbb{R}^{J+1}$ and $\mathbb{R}^{I+1}$, respectively.

Given these assumptions, transfers will be determined in equilibrium in each market as a function of market covariates and the proportions of types. We can then view transfers in each market as determined via a transfer function, which maps market characteristics to $\mathcal{T} \subseteq \mathbb{R}^{I \times J}$, the set of all possible transfer matrices Γ_t. Such a function may be denoted $\tilde{\gamma}_t$, where

$$\tilde{\gamma}_t : \mathcal{X}^{IJ} \times \mathcal{Z}^{IJ} \times \Delta^{I+J} \to \mathcal{T} \quad \text{such that} \quad \Gamma_t \equiv \tilde{\gamma}_t(X_t, Z_t, S_t),$$

with Δ^{I+J} a simplex in $\mathbb{R}^{I+J}$.

Assumptions 2.1 and 2.2 together ensure that the equilibrium proportions of match types and transfers only depend on market covariates X_t, Z_t and S_t. Were markets small (that is, without

Assumption 2.1), the match proportions and transfers would depend on the realizations of the idiosyncratic shocks. Were distributions of idiosyncratic shocks to vary across markets (that is, without Assumption 2.2(a)), matchings and transfers would vary with the idiosyncratic shock distributions even when markets are large. Hence, Assumptions 2.1 and 2.2 are precisely those required to ensure that the transfer functions $\tilde{\gamma}_t$ do not vary across markets t (henceforth, we drop the subscript t, $\Gamma_t = \tilde{\gamma}(X_t, Z_t, S_t)$ for all t).

We will now show that, in addition, X_t and Z_t enter the transfer function through linear indices. To this end, define

$$X_t\alpha \equiv (X_{11t}\alpha_1, X_{12t}\alpha_1, \cdots, X_{1Jt}\alpha_1, X_{21t}\alpha_2, X_{22t}\alpha_2, \cdots, X_{IJt}\alpha_I) \in \mathbb{R}^{IJ},$$

$$Z_t\beta \equiv (Z_{11t}\beta_1, Z_{21t}\beta_1, \cdots, Z_{I1t}\beta_1, Z_{12t}\beta_2, Z_{22t}\beta_2, \cdots, Z_{IJt}\beta_J) \in \mathbb{R}^{IJ},$$

to be the list of all possible indices in market t.

Proposition 2.3 *Under utility specification given by (2) and assumptions 2.1 and 2.2, there exists a transfer function* $\gamma : \mathbb{R}^{IJ} \times \mathbb{R}^{IJ} \to \mathcal{T} \subseteq \mathbb{R}^{I\times J}$ *such that*

$$\Gamma_t = \gamma(X_t\alpha, Z_t\beta, S_t)$$

for any market $t \in \mathbf{T}$.

The result follows from recalling that any equilibrium in a given market t maximizes total surplus in that market. Transfers that depend directly on market covariates X_t and Z_t, and not only on the indices $X_t\alpha$ and $Z_t\beta$ are not consistent with surplus maximization. The argument is made formally in the Appendix. Throughout, we will denote the $(i, j)^{th}$ element of $\gamma(X_t\alpha, Z_t\beta, S_t)$ by $\gamma_{ij}(X_t\alpha, Z_t\beta, S_t)$.

Proposition 2.3 is crucial in what follows. We use this result to map the observed match probabilities to the parameters $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_I)$ where $\alpha_i \in \mathbb{R}^{d_x}$ for each $i \in \mathbf{I}$, and $\beta = (\beta_1, \beta_2, \cdots, \beta_J)$ where $\beta_j \in \mathbb{R}^{d_z}$ for each $j \in \mathbf{J}$. Aggregating these across the T markets we are left with a conditional mean model, which links the observed data to model parameters α and β through an infinite-dimensional nuisance parameter (which arises since we do not make any parametric assumptions on the distribution of ε and ν, and since the transfer function is unobserved). This proposition is the key insight from this paper that links the matching model to a semiparametric multi-index model. Section 3 studies this link more carefully.

3 Reduced Form Model

Before we establish a link between the data and the parameters of interest, we first look at the match probabilities in a fixed market $t \in \mathbf{T}$, where $\mathbf{T}$ is a sample of markets. There are two separate discrete-choice problems in market t, one for side $\mathbf{a}$ and one for side $\mathbf{b}$. Define $|\mathbf{N}_{\mathbf{a}t}| + |\mathbf{N}_{\mathbf{b}t}| = N_t$. Recall that the proportions of types in the market is given by $S_t = (s_{\mathbf{a}1t}, \cdots, s_{\mathbf{a}It}, s_{\mathbf{b}1t}, \cdots, s_{\mathbf{b}Jt})$, where $s_{\mathbf{a}it}$ is the proportion of side $\mathbf{a}$ agents that are type i, and $s_{\mathbf{b}jt}$ is the proportion of side $\mathbf{b}$ agents that are type j.

Define

$$\mathbf{p}^{\mathbf{a}}_{ijt} \equiv \frac{\pi_{ijt}}{s_{\mathbf{a}it} N_t}$$

to be the match probability that type i agents on side $\mathbf{a}$ matches with type j, and

$$\mathbf{p}^{\mathbf{b}}_{ijt} \equiv \frac{\pi_{ijt}}{s_{\mathbf{b}jt} N_t}$$

to be the match probability that type j agents on side $\mathbf{b}$ that matches with type i, conditional on market covariates $X_t = X, Z_t = Z$. In the marriage example of Section 2.1, $s_{\mathbf{a}it}$ is the proportion of men that are type i and N_t is the total number of agents in the market. Therefore, $s_{\mathbf{a}it}N_t$ is the total number of type i men. Similarly, $s_{\mathbf{b}jt}N_t$ is total number of type j women. The above expressions for $\mathbf{p}^{\mathbf{a}}_{ijt}$ and $\mathbf{p}^{\mathbf{b}}_{ijt}$ transform the match outcome matrix in terms of conditional choice probabilities for each type on both sides.

Also denote the joint distribution of $\left\{(\varepsilon_{kj'} - \varepsilon_{kj}) : \mathbf{i}(k) = i,\ j' \in \mathbf{J} \cup \{0\}, \text{ and } j' \neq j\right\}$ by F_{ij}, and the joint distribution of $\{(\nu_{i'l} - \nu_{il}) : \mathbf{j}(l) = j,\ i' \in \mathbf{I} \cup \{0\}, \text{ and } i' \neq i\}$ by G_{ij}. Then for each type $i \in \mathbf{I}$ we have,

$$\mathbf{p}^{\mathbf{a}}_{ijt} = F_{ij}(X_{ij}\alpha_i - \tau_{ij}, \cdots, X_{ij}\alpha_i - X_{ij-1}\alpha_i - \tau_{ij} + \tau_{ij-1}, X_{ij}\alpha_i - X_{ij+1}\alpha_i - \tau_{ij} + \tau_{ij+1}, \cdots) \quad (5)$$

for all $j \in \mathbf{J} \cup \{0\}$, such that $\sum_{j=0}^{J} \mathbf{p}^{\mathbf{a}}_{ijt} = 1$. Similarly, for each type $j \in \mathbf{J}$ we have,

$$\mathbf{p}^{\mathbf{b}}_{ijt} = G_{ij}(Z_{ij}\beta_j + \tau_{ij}, \cdots, Z_{ij}\beta_j - Z_{i-1j}\beta_j + \tau_{ij} - \tau_{i-1j}, Z_{ij}\beta_j - Z_{i+1j}\beta_j + \tau_{ij} - \tau_{i+1j}, \cdots) \quad (6)$$

for all $i \in \mathbf{I} \cup \{0\}$, such that $\sum_{i=0}^{I} \mathbf{p}^{\mathbf{b}}_{ijt} = 1$.

Under the assumptions 2.1 and 2.2 we know that $\tau_{ij} = \gamma_{ij}(X_t\alpha, Z_t\beta, S_t)$. Therefore, (5) and (6) can just be written as $\tilde{H}^{\mathbf{a}}_{ij}(X_t\alpha, Z_t\beta, S_t)$ and $\tilde{H}^{\mathbf{b}}_{ij}(X_t\alpha, Z_t\beta, S_t)$, respectively. The following observation summarizes the main properties of the functions $\tilde{H}^{q}_{ij}$, for $q \in \{\mathbf{a}, \mathbf{b}\}$.

Observation 3.1 *For any* $q \in \{\mathbf{a}, \mathbf{b}\}$,

(a) $\tilde{H}^{q}_{ij}$ *does not depend on* t. *This is because* $F_{ij}, G_{ij}, \gamma_{ij}$ *are market invariant.*

(b) $\tilde{H}^{q}_{ij} : \mathbb{R}^{2IJ} \times \Delta^{I+J} \to [0,1]$.

(c) $\sum_{j=0}^{J} \tilde{H}^{\mathbf{a}}_{ij}(X_t\alpha, Z_t\beta, S_t) = 1$ *for all* $i \in \mathbf{I}$. *And,* $\sum_{i=0}^{I} \tilde{H}^{\mathbf{b}}_{ij}(X_t\alpha, Z_t\beta, S_t) = 1$ *for all* $j \in \mathbf{J}$.

(d) *There will be* $IJ + I$ *functions of the form* $\tilde{H}^{\mathbf{a}}_{ij}$, *and* $IJ + J$ *functions of the type* $\tilde{H}^{\mathbf{b}}_{ij}$.

In fact in every market $t \in \mathbf{T}$ we will have the following system of $2IJ + I + J$ equations,

$$
\begin{aligned}
\mathbf{p}^{\mathbf{a}}_{1jt} &= \tilde{H}^{\mathbf{a}}_{1j}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, j \in \mathbf{J} \cup \{0\}, &\quad \sum_{j=0}^{J} \mathbf{p}^{\mathbf{a}}_{1jt} = 1,\\
\mathbf{p}^{\mathbf{a}}_{2jt} &= \tilde{H}^{\mathbf{a}}_{2j}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, j \in \mathbf{J} \cup \{0\}, &\quad \sum_{j=0}^{J} \mathbf{p}^{\mathbf{a}}_{2jt} = 1,\\
&\vdots\\
\mathbf{p}^{\mathbf{a}}_{Ijt} &= \tilde{H}^{\mathbf{a}}_{Ij}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, j \in \mathbf{J} \cup \{0\}, &\quad \sum_{j=0}^{J} \mathbf{p}^{\mathbf{a}}_{Ijt} = 1,\\
\mathbf{p}^{\mathbf{b}}_{i1t} &= \tilde{H}^{\mathbf{b}}_{i1}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, i \in \mathbf{I} \cup \{0\}, &\quad \sum_{i=0}^{I} \mathbf{p}^{\mathbf{b}}_{i1t} = 1,\\
\mathbf{p}^{\mathbf{b}}_{i2t} &= \tilde{H}^{\mathbf{b}}_{i2}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, i \in \mathbf{I} \cup \{0\}, &\quad \sum_{i=0}^{I} \mathbf{p}^{\mathbf{b}}_{i2t} = 1,\\
&\vdots\\
\mathbf{p}^{\mathbf{b}}_{iJt} &= \tilde{H}^{\mathbf{b}}_{iJ}(X_t\alpha, Z_t\beta, S_t), &\quad \forall\, i \in \mathbf{I} \cup \{0\}, &\quad \sum_{i=0}^{I} \mathbf{p}^{\mathbf{b}}_{iJt} = 1.
\end{aligned}
\tag{7}
$$

The econometrician will observe $y^{\mathbf{a}}_{ijt}$, the sample proportion of type i agents on side $\mathbf{a}$ that matches with type j, and $y^{\mathbf{b}}_{ijt}$, the sample proportion of type j agents on side $\mathbf{b}$ that match with type i.[7] Define

$$Y_t = \begin{pmatrix} Y^{\mathbf{a}}_t \\ Y^{\mathbf{b}}_t \end{pmatrix},$$

where

$$Y^{\mathbf{a}}_t = (y^{\mathbf{a}}_{10t}, y^{\mathbf{a}}_{11t}, \cdots, y^{\mathbf{a}}_{1Jt}, \cdots, y^{\mathbf{a}}_{I0t}, {}^{\mathbf{a}}_{I1t}, \cdots, y^{\mathbf{a}}_{IJt}),$$

[7] Constructing $y^{\mathbf{a}}_{ijt}$ and $y^{\mathbf{b}}_{ijt}$, from a sample $\hat{N}_t$ of agents in market t. Let $\hat{\Pi}_t$ be the match allocation in the sample observed in market t, where its elements $\hat{\pi}_{ijt}$ denotes the observed number of (i,j) matches in the sample. Let $\hat{S}_t = (\hat{s}_{\mathbf{a}1t}, \cdots, \hat{s}_{\mathbf{a}It}, \hat{s}_{\mathbf{b}1t}, \cdots, \hat{s}_{\mathbf{b}Jt})$, be the observed distribution of types in the sample of $\hat{N}_t$ agents. That is, $\hat{s}_{\mathbf{a}it}$ is the observed proportion of side $\mathbf{a}$ agents that are type i, and $\hat{s}_{\mathbf{b}jt}$ is the observed proportion of side $\mathbf{b}$ agents that are type j, in the sample. Define

$$y^{\mathbf{a}}_{ijt} \equiv \frac{\hat{\pi}_{ijt}}{\hat{s}_{\mathbf{a}it}\hat{N}_t} \quad \text{and} \quad y^{\mathbf{b}}_{ijt} \equiv \frac{\hat{\pi}_{ijt}}{\hat{s}_{\mathbf{b}jt}\hat{N}_t}.$$

$$Y_t^{\mathbf{b}} = (y_{01t}^{\mathbf{b}}, y_{11t}^{\mathbf{b}}, \cdots, y_{I1t}^{\mathbf{b}}, \cdots, y_{0Jt}^{\mathbf{b}}, y_{1Jt}^{\mathbf{b}}, \cdots, y_{IJt}^{\mathbf{b}}),$$

are vectors of dimension $IJ + I$ and $IJ + J$, respectively. Therefore in every market $t \in \mathbf{T}$, the econometrician observes, $Y_t \in \mathcal{Y} \subseteq [0,1]^{2IJ+I+J}$, $X_t \in \mathcal{X}^{IJ}$, and $Z_t \in \mathcal{Z}^{IJ}$.

Let $e_{ijt}^{\mathbf{a}}$ and $e_{ijt}^{\mathbf{b}}$ be the projection error, i.e. it is defined to be the difference between the true and sample match proportions. Then we have,

$$y_{ijt}^{\mathbf{a}} = \mathbf{p}_{ijt}^{\mathbf{a}} + e_{ijt}^{\mathbf{a}},$$

$$y_{ijt}^{\mathbf{b}} = \mathbf{p}_{ijt}^{\mathbf{b}} + e_{ijt}^{\mathbf{b}},$$

such that

$$E\left(e_{ijt}^{\mathbf{a}} | X_t, Z_t, S_t\right) = E\left(e_{ijt}^{\mathbf{b}} | X_t, Z_t, S_t\right) = 0. \tag{8}$$

We can now write the system of equations (7) as,

$$E(Y_t | X_t, Z_t, S_t) = \tilde{H}(X_t\alpha, Z_t\beta, S_t) \tag{9}$$

where $\tilde{H}$ is vector of functions that does not change with t, of dimension $2IJ + I + J$. In particular, $\tilde{H} : \mathbb{R}^{2IJ} \times \Delta^{I+J} \to [0,1]^{2IJ+I+J}$. The components of $\tilde{H}$ are given by

$$E[y_{ijt}^q | X_t, Z_t, S_t] = \tilde{H}_{ij}^q(X_t\alpha, Z_t\beta, S_t),$$

for $q = \mathbf{a}, (i,j) \in \mathbf{I} \times (\mathbf{J} \cup \{0\})$, and $q = \mathbf{b}, (i,j) \in (\mathbf{I} \cup \{0\}) \times \mathbf{J}$. Further, integrating equation (9) over S_t, by law of iterated expectation we have that

$$E[Y_t | X_t, Z_t] = H(X_t\alpha, Z_t\beta), \tag{10}$$

where $H(X_t\alpha, Z_t\beta) = \int \tilde{H}(X_t\alpha, Z_t\beta, S_t) dF_S$, and F_S is the CDF of S_t.

Equation (10) gives the reduced form model. This maps the data $\{(Y_t, X_t, Z_t) : t \in \mathbf{T}\}$ to the unknown parameters of the model $(H, \{\alpha_i\}_{i \in \mathbf{I}}, \{\beta_j\}_{j \in \mathbf{J}})$. The goal of this paper is to identify the systematic preference parameters $\alpha_1, \alpha_2, \ldots, \alpha_I$ and $\beta_1, \beta_2, \ldots, \beta_J$ in the presence of the infinite-dimensional nuisance parameter H. This is a semiparametric multi-index model with multiple equations. Note that without proposition 2.3 we could not have written the conditional mean function in terms of the indices alone.

The Marriage Example *(continued from Section 2.1)* Recall that $X_{ijt} = (X_{ijt,educ}, X_{jt,inc})$ and $Z_{ijt} = (Z_{ijt,educ}, Z_{it,inc})$, for $(i,j) \in \{20,40\} \times \{20,40\}$. For instance, $X_{20,40,t} = (X_{20,40,t,educ}, X_{40,t,inc})$ where $X_{20,40,t,educ}$ is the absolute difference between expected education of 20-year-old men and expected education of 40-year-old women in market t, and $X_{40,t,inc}$ is the expected income of 40-

year-old men in market t. Here, we will observe the proportions of 20-year-old men matching 20-year-old women ($y^{\mathbf{a}}_{20,20,t}$), 40-year-old women ($y^{\mathbf{a}}_{20,40,t}$), and proportions remaining unmatched ($y^{\mathbf{a}}_{20,0,t}$). Similarly, we observe matching proportions for 40-year-old men ($y^{\mathbf{a}}_{40,20,t}$,$y^{\mathbf{a}}_{40,40,t}$,$y^{\mathbf{a}}_{40,0,t}$), 20-year-old women ($y^{\mathbf{b}}_{20,20,t}$,$y^{\mathbf{b}}_{40,20,t}$,$y^{\mathbf{b}}_{0,20,t}$), and 40-year-old women ($y^{\mathbf{b}}_{20,40,t}$,$y^{\mathbf{b}}_{40,40,t}$,$y^{\mathbf{b}}_{0,40,t}$). Therefore, the outcome vector Y_t is given by,

$$Y_t = (y^{\mathbf{a}}_{20,0,t}, y^{\mathbf{a}}_{20,20,t}, y^{\mathbf{a}}_{20,40,t}, y^{\mathbf{a}}_{40,0,t}, y^{\mathbf{a}}_{40,20,t}, y^{\mathbf{a}}_{40,40,t}, y^{\mathbf{b}}_{0,20,t}, y^{\mathbf{b}}_{20,20,t}, y^{\mathbf{b}}_{40,20,t}, y^{\mathbf{b}}_{0,40,t}, y^{\mathbf{b}}_{20,40,t}, y^{\mathbf{b}}_{40,40,t})'.$$

Any component of Y_t (say, $y^{\mathbf{b}}_{40,20,t}$) can be written as

$$\begin{aligned} E(y^{\mathbf{b}}_{40,20,t}|X_t, Z_t) \quad = \quad & H^{\mathbf{b}}_{40,20}(X_{20,20,t}\alpha_{20}, X_{20,40,t}\alpha_{20}, X_{40,20,t}\alpha_{40}, X_{40,40,t}\alpha_{40}, \\ & Z_{20,20,t}\beta_{20}, Z_{20,40,t}\beta_{40}, Z_{40,20,t}\beta_{20}, Z_{40,40,t}\beta_{40}) \end{aligned}$$

where $H^{\mathbf{b}}_{40,20}$ is an unknown function with eight arguments. There are twelve such equations (one for each component of Y_t) with cross-equation restrictions such as, $H^{\mathbf{b}}_{20,20} + H^{\mathbf{b}}_{40,20} + H^{\mathbf{b}}_{0,20} = 1$.

4 Identification

In the previous section, we reduced the matching model in T markets to the following semiparametric multi-index model,

$$E(Y_t|X_t, Z_t) = H(X_t\alpha, Z_t\beta), \quad t \in \mathbf{T}.$$

Multi-index models have been studied in detail in Ichimura and Lee (1991) when there is only one equation. Here, we have $2IJ + I + J$ with cross-equation restrictions. In this section, we study the identification of preference parameters $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_I) \in \mathbb{R}^{d_x} \times \mathbb{R}^I$ and $\beta = (\beta_1, \beta_2, \cdots, \beta_J) \in \mathbb{R}^{d_z} \times \mathbb{R}^J$ under conditions that are similar to Ichimura and Lee (1991). The main difference is in the exclusion restrictions we impose which are specific to the application here.

Let $\mathcal{H}$ denote the parameter space of H. For any vector function $H \in \mathcal{H}$, $H : \mathbb{R}^{2IJ} \to [0,1]^{2IJ+I+J}$. In addition, $H(\cdot)$ will contain J distinct subvectors of dimension $I+1$, each of which will sum to one, and also I distinct subvectors of dimension $J+1$, each of which will also sum to 1. Despite these restrictions, $H \in \mathcal{H}$ is an unknown vector function and thus there exists an indeterminacy between the unknown vector function H and the parameters of interest α and β. That is, without imposing further restrictions on the parameters and the covariates, it is not possible to identify $(H, \alpha, \beta) \in \mathcal{H} \times \mathbb{R}^{Id_x} \times \mathbb{R}^{Jd_z}$. The following definitions are introduced to formalize the idea of non-identification in Lemma 4.1.

DEFINITION: Two distinct parameter vectors (α, β) and (α^*, β^*) are said to be *observationally equivalent* ($\overset{o.e.}{\sim}$) if there exist $H, H^* \in \mathcal{H}$ such that $E(Y|X, Z) = H(X\alpha, Z\beta) = H^*(X\alpha^*, Z\beta^*)$.

DEFINITION: (α,β) is said to be *identified* if there does not exist $(\alpha^*,\beta^*) \neq (\alpha,\beta)$ such that $(\alpha^*,\beta^*) \overset{o.e.}{\sim} (\alpha,\beta)$.

We can write $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_I) = (\alpha_i,\alpha_{-i})$ for any $i \in \mathbf{I}$ where $\alpha_{-i} = (\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_I) \in \mathbb{R}^{d_x} \times \mathbb{R}^{I-1}$. Similarly, we can write $\beta = (\beta_1,\beta_2,\ldots,\beta_J) = (\beta_j,\beta_{-j})$ for any $j \in \mathbf{J}$ where $\beta_{-j} = (\beta_1,\ldots,\beta_{j-1},\beta_{j+1},\ldots,\beta_J) \in \mathbb{R}^{d_z} \times \mathbb{R}^{J-1}$.

Lemma 4.1 *In the multiple index model given by (10) we have the following:*

(a) For any $c \neq 0$, given $i \in \mathbf{I}$ and $j \in \mathbf{J}$, let $\alpha^* = (c\alpha_i,\alpha_{-i})$ and $\beta^* = (c\beta_j,\beta_{-j})$. Then, $(\alpha^*,\beta)\overset{o.e.}{\sim}(\alpha,\beta^*)\overset{o.e.}{\sim}(\alpha,\beta)$.

(b) If $X_{ij} = X_{i'j}$ for $i \neq i'$, then $(\alpha,\beta)\overset{o.e.}{\sim}(\alpha^*,\beta)$ where $\alpha^* = (\alpha_i - \alpha_{i'},\alpha_{-i})$. Similarly, if $Z_{ij} = Z_{ij'}$ for $j \neq j'$, then $(\alpha,\beta)\overset{o.e.}{\sim}(\alpha,\beta^*)$ where $\beta^* = (\beta_j - \beta_{j'},\beta_{-j})$.

(c) If X_{ij} is a subvector of $Z_{i'j'}$, then $(\alpha,\beta)\overset{o.e.}{\sim}(\alpha,\beta^*)$ where $\beta^* = (\beta_{j'} - \mu_1\alpha,\beta_{-j'})$, and μ_1 is a $d_z \times d_x$ selection matrix such that $X_{ij} = Z_{i'j'}\mu_1$. Similarly, if Z_{ij} is a subvector of $X_{i'j'}$, then $(\alpha,\beta)\overset{o.e.}{\sim}(\alpha^*,\beta)$ where $\alpha^* = (\alpha_{i'} - \mu_2\beta,\alpha_{-i'})$, and μ_2 is a $d_x \times d_z$ selection matrix such that $Z_{ij} = X_{i'j'}\mu_2$.

Lemma 4.1 suggests some necessary conditions for identification of α and β. Part (a) implies that we need to restrict the parameter space of α and β in such a way that scaling is not possible for any of the coefficients of the linear indices. Here it is convenient to assume

$$\alpha_i = (1,\tilde{\alpha}_i),\ \forall\, i \in \mathbf{I} \quad \text{and} \quad \beta_j = (1,\tilde{\beta}_j),\ \forall\, j \in \mathbf{J}.$$

Parts (b) and (c) suggest that we will need some exclusion restrictions. If X_{ij} is the same random variable as $X_{i'j}$, then α_i cannot be distinguished from $\alpha_{i'}$. Therefore, there needs to be at least one component of X_{ij} that is not included in $X_{i'j}$ and vice-versa. This implies we cannot have X_{ij} invariant with respect to i. That is, we cannot have $X_{ij} = X_{i'j} = X_j$. Thus, we need at least one variable in the vector X_{ij} (say $X_{ij}^{(1)}$) that varies with match type (i,j). However, if $\alpha_1 = \alpha_2 = \cdots = \alpha_I$, then we do not need this exclusion. Similarly, Z_{ij} cannot be invariant with respect to j, so we can distinguish between β_j and $\beta_{j'}$. Finally, we also need exclusion restrictions so that we can distinguish between the α's and β's. The following set of exclusion restrictions are sufficient to rule out these non-identifiable cases.

Assumption 4.2 *(a) Each X_{ij}, $(i,j) \in \mathbf{I}\times\mathbf{J}$, contains a continuous explanatory variable $X_{ij}^{(1)}$ with a non-zero coefficient normalized to one. That is, $X_{ij}\alpha_i = X_{ij}^{(1)} + X_{ij}^{(2)}\tilde{\alpha}_i$. Further, $X_{ij}^{(1)} \neq X_{i'j}^{(1)}$ for any $i \neq i'$.*

(b) *Each* Z_{ij}, $(i,j) \in \mathbf{I} \times \mathbf{J}$, *contains a continuous explanatory variable* $Z_{ij}^{(1)}$ *with a non-zero coefficient normalized to one. That is,* $Z_{ij}\beta_j = Z_{ij}^{(1)} + Z_{ij}^{(2)}\tilde{\beta}_j$. *Further,* $Z_{ij}^{(1)} \neq Z_{ij'}^{(1)}$ *for any* $j \neq j'$.

(c) *For every* $(i,j),(i',j') \in \mathbf{I} \times \mathbf{J}$, $X_{ij}^{(2)}$ *is not contained in the covariate vector* $Z_{i'j'}^{(2)}$ *(and, vice-versa).*

First note that we only require the existence of one variable that is continuously distributed in each index. This assumption is necessary because, if all the regressors were discrete, then the model has no information. Bierens and Hartog (1988) shows this result for a single-index model. However, we do not require a large support assumption for identification of α and β.

Assumption 4.2 helps eliminate the non-identifiable cases. We can understand this in the context of the marriage example in Section 2.1. Here, agents care about having similar expected education levels to their spouse. This is captured by the random variable $X_{ij}^{(1)} = X_{ij,educ} \equiv |X_{i,educ} - X_{j,educ}|$, which is a different random variable for each pair (i,j). Therefore, in each index included in the list $X_t\alpha$, $X_{ij}^{(1)}$ will be the excluded random variable. That is, $X_{ij}^{(1)}$ and $X_{i'j}^{(1)}$ are two separate random variables that act as instruments for distinguishing between α_i and $\alpha_{i'}$, where $i \neq i'$. For this same reason, we can set $Z_{ij}^{(1)} = X_{ij}^{(1)}$, then the excluded variables $X_{ij}^{(1)}$ and $X_{ij'}^{(1)}$ act as instruments that distinguish between β_j and $\beta_{j'}$, where $j \neq j'$. Thus, including a covariate that is different for each possible pair (i,j) satisfies Assumptions 4.2(a)-(b). Also, we include $X_{j,inc}$ in $X_{ij}^{(2)}$ and $Z_{i,inc}$ in $Z_{ij}^{(2)}$. Note that $X_{j,inc}$ and $Z_{i,inc}$ are $I+J$ different random variables, and thus Assumption 4.2(c) is satisfied. Here, $X_{j,inc}$ and $Z_{i,inc}$ act as instruments to distinguish between the α's and the β's.

We further impose some smoothness and non-linearity assumptions on H which will be sufficient for identification of the finite dimensional parameters α and β.

Assumption 4.3 (a) *Each* H_{ij}^q, *where* $q = \mathbf{a}, (i,j) \in \mathbf{I} \times (\mathbf{J} \cup \{0\})$, *or* $q = \mathbf{b}, (i,j) \in (\mathbf{I} \cup \{0\}) \times \mathbf{J}$, *is differentiable with respect to the continuous covariates.*

(b) *Let* $H_{ij,r}^q$ *denote the partial derivative of* H_{ij}^q *with respect to the* r^{th} *argument. Then, for each* $q = \mathbf{a}, (i,j) \in \mathbf{I} \times (\mathbf{J} \cup \{0\})$ *and* $q = \mathbf{b}, (i,j) \in (\mathbf{I} \cup \{0\}) \times \mathbf{J}$, *the derivative functions* $H_{ij,r}^q$, $r = 1, \cdots, 2IJ$ *are linearly independent with probability one on* $\mathcal{X}^{IJ} \times \mathcal{Z}^{IJ}$.

The first part of Assumption 4.3 is simply a smoothness restriction on each component function of vector function H. The second part of this assumption imposes a mild restriction on the shape of the component functions of H. It requires the partial derivative of each component of H to be linearly independent. A sufficient condition for this to hold is for the generalized Wronskian for derivative functions $H_{ij,r}^q$, $r = 1, \cdots, 2IJ$ to not be zero.[8] A further sufficient condition for the generalized

[8] Generalized Wronskians are determinant of the Jacobian matrix of the derivative function. To check this condition we will need the derivative functions $H_{ij,r}^q$ to be differentiable with respect to the continuous covariates.

Wronskian to not be zero is that H^q_{ij} is non-linear in each of its arguments. This does not seem so strong since H^q_{ij} is made up of joint distribution functions F_{ij} or G_{ij} which are typically non-linear.

Our identification strategy is such that Assumption 4.3 is only sufficient to identify components of α_i and β_j that are coefficients to continuous covariates. However, we allow $X^{(2)}_{ij}$ and $Z^{(2)}_{ij}$ to contain discrete random variables.[9] Let ρ^x_{ij} be a vector of discrete covariates contained in $X^{(2)}_{ij}$, and ρ^z_{ij} be a vector of discrete covariates contained in $Z^{(2)}_{ij}$. Denote the subvector of α_i that is the coefficient of ρ^x_{ij} by $\tilde{\tilde{\alpha}}_i$, and the subvector of β_j that is the coefficient of ρ^z_{ij} by $\tilde{\tilde{\beta}}_j$. Define $\tilde{\tilde{\alpha}} = (\tilde{\tilde{\alpha}}_1, \dots, \tilde{\tilde{\alpha}}_I)$ and $\tilde{\tilde{\beta}} = (\tilde{\tilde{\beta}}_1, \dots, \tilde{\tilde{\beta}}_J)$. Then we need the following additional assumption:

Assumption 4.4 *Let $\rho_{ijt} \in \Omega$ be a vector consisting of distinct components of ρ^x_{ij} and ρ^z_{ij}. Then, there exists two points $\rho = (\rho^x_{11}, ...\rho^x_{ij}, \dots, \rho^z_{11}, \dots \rho^z_{ij}, \dots)$ and $\bar{\rho} = (\bar{\rho}^x_{11}, ...\bar{\rho}^x_{ij}, \dots, \bar{\rho}^z_{11}, \dots \bar{\rho}^z_{ij}, \dots)$ in Ω^{2IJ} such that if $\tilde{\tilde{\alpha}}_0 \neq \tilde{\tilde{\alpha}}$ and $\tilde{\tilde{\beta}}_0 \neq \tilde{\tilde{\beta}}$, then*

$$H(\cdots, X_{ijt}\alpha_i + \rho^x_{ijt}(\tilde{\tilde{\alpha}}_{i0} - \tilde{\tilde{\alpha}}_i), \cdots, Z_{ijt}\beta_j + \rho^z_{ijt}(\tilde{\tilde{\beta}}_{j0} - \tilde{\tilde{\beta}}_j), \cdots)$$
$$\neq H(\cdots, X_{ijt}\alpha_i + \bar{\rho}^x_{ijt}(\tilde{\tilde{\alpha}}_{i0} - \tilde{\tilde{\alpha}}_i), \cdots, Z_{ijt}\beta_j + \bar{\rho}^z_{ijt}(\tilde{\tilde{\beta}}_{j0} - \tilde{\tilde{\beta}}_j), \cdots).$$

That is, to identify coefficients corresponding to discrete regressors, we need at least two distinct points in its support at which the function can be distinguished.

Proposition 4.5 *Let Assumptions 2.1-2.2 and 4.2-4.4 hold. Then, α and β are identified (up to a multiplicative scalar) in the multi-index model given by (10).*

We have point identification of parameter vectors α and β. However, we have said nothing so far about the identification of the link function H. Note that $H : \mathbb{R}^{2IJ} \to [0,1]^{2IJ+I+J}$. Without a large (i.e., unbounded) support assumption it may not be possible to point identify the unknown vector of link functions. Instead, under Assumption 4.2(a), H will be partially identified. Let $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_{-1} \subseteq \mathbb{R} \times \mathbb{R}^{d_x - 1}$, where $\mathcal{X}_1$ denotes the support of the excluded regressor $X^{(1)}_{ij}$, and let $\mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_{-1} \subseteq \mathbb{R} \times \mathbb{R}^{d_z - 1}$, where $\mathcal{Z}_1$ denotes the support of the excluded regressor $Z^{(1)}_{ij}$, for every $(i,j) \in \mathbf{I} \times \mathbf{J}$. The following proposition discusses identification of H.

Proposition 4.6 *Let Assumptions 2.1-2.2 and 4.2-4.4 hold.*

(a) If $\mathcal{X}_1 = \mathbb{R}$ and $\mathcal{Z}_1 = \mathbb{R}$, then H is point identified.

(b) If $\mathcal{X}$ is a compact subset of $\mathbb{R}^{d_x}$, or $\mathcal{Z}$ is a compact subset of and $\mathbb{R}^{d_z}$, then $H \in \mathcal{H}$ is partially identified. Further, the identified set, $\mathcal{H}_I$, will be a strict subset of $\mathcal{H}$.

H is not directly the object of interest in this paper. However, depending on the estimation method, the point or partial identification of H may be relevant for doing inference on α and β.

[9] In the marriage example in Section 2.1, we could have taken$X_{j,inc}$ as the mode of income instead of the expected income, for type j women (where income has discrete support {high=3, medium=2, low=1},say). In this case for every $j \in \mathbf{J}$, $X_{j,inc}$ is a discrete random variable.

5 Conclusion

This paper considers the identification problem in a one-to-one matching model with transferable utilities. This framework builds on Choo and Siow (2006), where the objective is to identify systematic preferences from the knowledge of match patterns observed over some characteristics. We extend this framework in two salient ways. First, we explicitly model the systematic preferences of agents by introducing covariates through linear indices. Introducing covariates in preferences aids us in understanding the role of the different covariates in determining a match outcome. Second, we consider match patterns only over a subset of observed characteristics (which we call types). The systematic preferences for each match type are then described by the remaining observed characteristics. This allows us to discern preferences based on several observed agent characteristics, without increasing the dimension of the match outcome matrix, which describes matching patterns over types. We also allow for unobserved heterogeneity in the agent preferences, but make no distributional assumptions over it. Instead, we consider a sample from multiple markets. With such preferences and some additional assumptions, we show that transfers depend only on systematic preferences. Using this result on transfers, the matching model is reduced to a semiparametric multi-index model. We then study identification of the relevant parameters in this reduced form.

Several issues remain for future research. The systematic preferences are modeled with covariates, which are aggregated at the level of types. This imposes a qualitative restriction on agent preferences that every match partner of same type is exchangeable. The question then is if we can relax exchangeability in these models. Another question is how much one can learn from considering just one market when knowledge of the distribution of the taste shocks is not assumed.

In principal, semiparametric multi-index models can be estimated using sieve minimum distance estimator of Ai and Chen (2003), or a semiparametric kernel estimator of Ichimura and Lee (1991). However, the high dimensionality of the functions in H prevents these estimators from performing well in finite sample. Hristache et al. (2001) provides a method for dimension reduction in multiple-index models, thus selecting smaller number of relevant indices. One can possibly construct estimators in the family of average derivative estimators extending methods of Hristache, Juditsky, and Spokoiny (2001), which is beyond the scope of this paper.

References

[1] Agarwal, N. and W. Diamond (2017): "Latent Indices In Assortative Matching Models," *Quantitative Economics*,8(3), 685–728.

[2] Ai, C. and X. Chen (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71 (6), 1795-1843.

[3] Banerjee, A., E. Duflo, M. Ghatak, and J. Lafortune (2013): "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, 5 (2), 33-72.

[4] Becker, G. (1973): "A Theory of Marriage, part I," *Journal of Political Economy*, 81 (4), 813-846.

[5] Becker, G. (1974): "A Theory of Marriage, part II," *Journal of Political Economy*, 82 (2), S11-26.

[6] Bierens, H.J. and J. Hartog (1988): "Non-Linear Regression with Discrete Explanatory Variables, with an Application to the Earnings Function," *Journal of Econometrics*, 38, 269-299.

[7] Botticini, M., and A. Siow (2008): "Are there Increasing Returns in Marriage Markets?," University of Toronto Department of Economics Working Paper 333.

[8] Chiappori, P.-A., S. Oreffice, and C. Quintana-Domeque (2012): "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market," *Journal of Political Economy*, 120 (4), 659-695.

[9] Chiappori, P.-A., and B. Salanié (2016): "The Econometrics of Matching Models," *Journal of Economic Literature, 54(3), 832-861.*

[10] Chiappori, P.-A., B. Salanié, and Y. Weiss (2017): "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review*, 107 (8), 2109-67.

[11] Choo, E., and A. Siow (2006): "Who Marries Whom and Why?," *Journal of Political Economy*, 114, 175-201.

[12] Dagsvik, J.K. (2000): "Aggregation in Matching Markets," *International Economic Review*, 41 (February), 27–57.

[13] Dupuy, A. and A. Galichon (2014): "Personality Traits and the Marriage Market," *Journal of Political Economy*, 122 (6), 1271-319.

[14] Echenique, F., S. Lee, M. Shum, and M.B. Yenmez (2013): "The Revealed Preference Theory of Stable and Extremal Stable Matchings," *Econometrica*, 81, 153-171.

[15] Fox, J.T. (2009): “Structural Empirical Work Using Matching Models,” in S.N. Durlauf and L.E. Blume (Eds.), *New Palgrave Dictionary of Economics* (Online ed.)

[16] Fox, J.T. (2010): “Identification in Matching Games,” *Quantitative Economics*, 1, 203-254.

[17] Fox, J.T., D. Hsu, and C. Yang (2017): “Unobserved Heterogeneity in Matching Games,” *Journal of Political Economy*, forthcoming.

[18] Gabaix, X. and A. Landier (2008): “Why has CEO pay increased so much?,” *The Quarterly Journal of Economics*, 123(1), 49-100.

[19] Galichon, A., and B. Salanié (2015): “Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models,” Columbia University Discussion Paper 1415-02.

[20] Galichon, A., S.D. Kominers and S. Weber (2016): “Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility,” Unpublished manuscript.

[21] Graham, B. (2011): “Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers,” in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. Jackson, Elsevier.

[22] Graham, B. (2013a): “ Errata in ’Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers’,” Unpublished manuscript.

[23] Graham, B. (2013b): “Comparative static and computational methods for an empirical one-to-one transferable utility matching model,” *Advances in Econometrics* 31: 153 - 181.

[24] Gretsky, N.E., J.M. Ostroy, and W.R. Zame (1992): “The nonatomic assignment model,” *Economic Theory*, 2, 103–127.

[25] Gretsky, N.E., J.M. Ostroy, and W.R. Zame (1999): “Perfect competition in the continuous assignment model,” *Journal of Economic Theory*, 88, 60–118.

[26] Hristache, M., A. Juditsky, and V. Spokoiny (2001): “Direct estimation of the index coefficients in a single-index model,” *Annals of Statistics*, 29, 1-32.

[27] Hristache, M., A. Juditsky, J. Polzehl, and V. Spokoiny (2001): “Structure Adaptive Approach for Dimension Reduction,” *Annals of Statistics*, 29, 1537-1566.

[28] Ichimura, H. and Lee, L.-F. (1991): “Semiparametric Least Squares Estimation of multiple index models: Single equation estimation,” in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, ed. by W.A. Barnett, J. Powell, and G. Tauchen.

[29] Koopmans, T.C. and M. Beckmann (1957): "Assignment Problems and the Location of Economic Activities," *Econometrica*, 25 (1), 53–76.

[30] Menzel, K. (2015): "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 83 (3), 897-941.

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

[32] Salanié, B. (2015): "Identification in Separable Matching with Observed Transfers," *Columbia University mimeo.*

[33] Shapley, L., and M. Shubik (1972): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111-130.

[34] Siow, A., and E. Choo (2006): "Estimating a Marriage Matching Model with Spillover Effects," *Demography*, 43(3), 463-490.

[35] Tervio, M. (2008): "The difference that CEOs make: An assignment model approach," *American Economic Review*, 98, 642-668.

Appendix

Proof of Proposition 2.3:

The proof builds on the observation that, if t' and t'' are distinct markets with $X_{t'}\alpha = X_{t''}\alpha$ and $Z_{t'}\beta = Z_{t''}\beta$ and $S_{t'} = S_{t''}$, then the equilibrium matching must be the same in each market (up to values of the preference shocks that have zero measure). This follows because the equilibrium matching in each market is efficient, and because efficient matchings are uniquely determined (up to values of the preference shocks that have zero measure; see, e.g., Gretsky, Ostroy and Zame (1999)). Given that the equilibrium matchings are identical in both markets, we then verify that equilibrium transfers must be identical.

Consider a transfer function

$$\bar{\gamma} : \mathcal{X}^{IJ} \times \mathbb{R}^{IJ} \times \mathcal{Z}^{IJ} \times \mathbb{R}^{IJ} \times \Delta^{I+J} \to \mathcal{T} \subseteq \mathbb{R}^{I\times J},$$

such that the equilibrium transfers Γ_t in each market t are given by the matrix $\bar{\gamma}(X_t, X_t\alpha, Z_t, Z_t\beta, S_t)$, with $(i,j)^{th}$ element $\bar{\gamma}_{ij}(X_t, X_t\alpha, Z_t, Z_t\beta, S_t)$. If the result in the proposition fails to hold, then there exist distinct markets t' and t'' with $X_{t'}\alpha = X_{t''}\alpha$, $Z_{t'}\beta = Z_{t''}\beta$ and $S_{t'} = S_{t''}$, but with $\bar{\gamma}_{ij}(X_{t'}, X_{t'}\alpha, Z_{t'}, Z_{t'}\beta, S_{t'}) \neq \bar{\gamma}_{ij}(X_{t''}, X_{t''}\alpha, Z_{t''}, Z_{t''}\beta, S_{t''})$ for some (i,j).

Suppose, without loss of generality, that

$$\bar{\gamma}_{ij}(X_{t'}, X_{t'}\alpha, Z_{t'}, Z_{t'}\beta, S_{t'}) < \bar{\gamma}_{ij}(X_{t''}, X_{t''}\alpha, Z_{t''}, Z_{t''}\beta, S_{t''}).$$

By Assumption 2.2(b), there exists a positive measure of shocks $\varepsilon_{k't'}$, $k' \in \mathbf{N}_{\mathbf{a}t'}$ with $\mathbf{i}(k') = i$, such that

$$X_{ijt'}\alpha_i + \bar{\varepsilon}_{k'jt'} - \bar{\gamma}_{ij}(X_{t'}, X_{t'}\alpha, Z_{t'}, Z_{t'}\beta, S_{t'}) > X_{ij't'}\alpha_i + \bar{\varepsilon}_{k'j't'} - \bar{\gamma}_{ij'}(X_{t'}, X_{t'}\alpha, Z_{t'}, Z_{t'}\beta, S_{t'})$$

for all $j' \in \mathbf{J} \setminus \{j\}$ and

$$X_{ijt'}\alpha_i + \bar{\varepsilon}_{k'jt'} - \bar{\gamma}_{ij}(X_{t'}, X_{t'}\alpha, Z_{t'}, Z_{t'}\beta, S_{t'}) > \bar{\varepsilon}_{k'0t'},$$

implying that the agent k' in market t' optimally chooses to match with j, and yet

$$\bar{\varepsilon}_{k'0t'} > X_{ij't'}\alpha_i + \bar{\varepsilon}_{k'j't'} - \bar{\gamma}_{ij'}(X_{t''}, X_{t''}\alpha, Z_{t''}, Z_{t''}\beta, S_{t''}),$$

for all $j' \in \mathbf{J}$, implying that the agent k' would prefer to remain unmatched given the transfers that prevail in market t'' (for this, it is enough to consider values $\bar{\varepsilon}_{k'j't'}$, for $j' \neq j$, that are sufficiently small). Hence, if $k'' \in \mathbf{N}_{\mathbf{a}t''}$ is an agent in market t'' with $\varepsilon_{k''t''} = \varepsilon_{k't'}$, we have

$$\bar{\varepsilon}_{k''0t''} > X_{ij't''}\alpha_i + \bar{\varepsilon}_{k'j't''} - \bar{\gamma}_{ij'}(X_{t''}, X_{t''}\alpha, Z_{t''}, Z_{t''}\beta, S_{t''}),$$

implying that the agent k'' prefers to remain unmatched in market t''. Since this holds for a positive measure of shocks, the match allocations are different for a positive measure of agents in the two markets t' and t'', contradicting that the (essentially) unique allocations in both sides coincide. ■

Proof of Lemma 4.1:

For Part (a) of the lemma, let the true parameters that satisfy equation (10) be given by $(H, \alpha, \beta) \in \mathcal{H} \times \mathbb{R}^{Id_x} \times \mathbb{R}^{Jd_z}$. Consider $\tilde{\alpha} = (\alpha_1, \cdots, \alpha_{i-1}, c\alpha_i, \alpha_{i+1}, \cdots, \alpha_I)$ for some $c \neq 0$. Let $X_{ijt}\alpha_i, j = 0, 1, 2, \cdots, J$ be the first $J+1$ arguments of the vector function H. Also split the first $IJ+I$ components of H into distinct subvector of $J+1$ dimension $(H_1^{\mathbf{a}}, H_2^{\mathbf{a}}, \cdots, H_I^{\mathbf{a}})$ where $H_i^{\mathbf{a}} = (h_{i0}^{\mathbf{a}}, h_{i1}^{\mathbf{a}}, ..., h_{iJ}^{\mathbf{a}})$. And, split the remaining $IJ + J$ components of H into distinct subvector of $I + 1$ dimension $(H_1^{\mathbf{b}}, H_2^{\mathbf{b}}, \cdots, H_J^{\mathbf{b}})$ where $H_j^{\mathbf{b}} = (h_{0j}^{\mathbf{b}}, h_{1j}^{\mathbf{b}}, ..., h_{Ij}^{\mathbf{b}})$.

Define a vector function $\tilde{H}$ such that, in $\tilde{H}_i^{\mathbf{a}}$ set $\tilde{h}_{i0}^{\mathbf{a}} = 1 - \sum_{j=1}^J \tilde{h}_{ij}^{\mathbf{a}}$ for all $i \in \mathbf{I}$; in $\tilde{H}_j^{\mathbf{b}}$ set $\tilde{h}_{0j}^{\mathbf{b}} = 1 - \sum_{i=1}^I \tilde{h}_{ij}^{\mathbf{b}}$ for all $j \in \mathbf{J}$; and $\tilde{h}_{ij}^q(\cdot, \cdot, \cdots) = h_{ij}^q\left(\frac{1}{c}\cdot, \cdot, \cdots\right)$, where only the first $J+1$ arguments are scaled by $\frac{1}{c}$, where $q \in \{\mathbf{a}, \mathbf{b}\}$ and $(i, j) \in \mathbf{I} \times \mathbf{J}$. Then, $(\tilde{H}, \beta, \tilde{\alpha}) \in \mathcal{H} \times \mathbb{R}^{Id_x} \times \mathbb{R}^{Jd_z}$. By construction we will have that $\tilde{H}(X_t\tilde{\alpha}, Z_t\beta) = H(X_t\alpha, Z_t\beta)$. Therefore the data cannot distinguish between α and $c\alpha$, and therefore we can only identify α up to a scaled normalization. Similarly we can scale components of β.

The arguments are identical for both (b) and (c). So we will just show it for part (c). Without loss of generality, suppose for every $(i, j) \in \mathbf{M}$, $X_{ijt} = Z_{ijt}\mu$ where μ is a $d_z \times d_x$ dimensional selection matrix. Let the true parameters that satisfy equation (10) be given by (H, α, β). Let $\tilde{\beta} = \beta - \mu\alpha$. Construct a vector function $\tilde{H}$ as follows:

In $\tilde{H}_i^{\mathbf{a}}$ set $\tilde{h}_{i0}^{\mathbf{a}} = 1 - \sum_{j=1}^J \tilde{h}_{ij}^{\mathbf{a}}$ for all $i \in \mathbf{I}$; and in $\tilde{H}_j^{\mathbf{b}}$ set $\tilde{h}_{0j}^{\mathbf{b}} = 1 - \sum_{i=1}^I \tilde{h}_{ij}^{\mathbf{b}}$ for all $j \in \mathbf{J}$, where

$$\tilde{h}_{i'j'}^q(\cdots, X_{ijt}\alpha_i, Z_{ijt}\tilde{\beta}_j, \cdots) = \tilde{h}_{i'j'}^q(\cdots, X_{ijt}\alpha_i, Z_{ijt}\beta_j - X_{ijt}\alpha_i, \cdots) = h_{i'j'}^q(\cdots, X_{ijt}\alpha_i, Z_{ijt}\beta_j, \cdots),$$

for any $q \in \{\mathbf{a}, \mathbf{b}\}$ and $(i', j') \in \mathbf{I} \times \mathbf{J}$. Therefore the data cannot distinguish between elements of β and $\tilde{\beta}$ that correspond to the coefficients of X_{ijt}. ■

Proof of Proposition 4.5:

Suppose there exists (α_0, β_0) and (α, β) such that

$$E[Y_t | X_t\alpha_0, Z_t\beta_0] = E[Y_t | X_t\alpha, Z_t\beta] = H(X_t\alpha_0, Z_t\beta_0) \tag{11}$$

almost everywhere in the support of (X_t, Z_t). By Assumption 4.2 let $\alpha_i = (1, \tilde{\alpha}_i)$ and $\beta_j = (1, \tilde{\beta}_j)$. Further we can write $X_{ijt}\alpha_i = X_{ijt}^{(1)} + X_{ijt}^{(2)}\tilde{\alpha}_i$ and $Z_{ijt}\beta_j = Z_{ijt}^{(1)} + Z_{ijt}^{(2)}\tilde{\beta}_j$ for all $(i,j) \in \mathbf{I} \times \mathbf{J}$. Let $X_{ijt}^{(2)} \in \mathcal{X}_{-1} \subseteq \mathbb{R}^{d_x-1}$ and $Z_{ijt}^{(2)} \in \mathcal{Z}_{-1} \subseteq \mathbb{R}^{d_z-1}$. Let $w_{ijt} \in W \subseteq \mathbb{R}^{d_w}$ be a vector of *distinct* elements of $X_{ijt}^{(2)}$ and $Z_{ijt}^{(2)}$. Assumption 4.2(d) implies $1 + \max(d_x - 1, d_z - 1) \leq d_w \leq d_x + d_z - 2$. Let $X_t^{(1)} \equiv (X_{11t}^{(1)}, \cdots, X_{1Jt}^{(1)}, \cdots, X_{IJt}^{(1)})$ be an IJ dimensional vector, $Z_t^{(1)} \equiv (Z_{11t}^{(1)}, \cdots, Z_{I1t}^{(1)}, \cdots, Z_{IJt}^{(1)})$ be an IJ dimensional vector, and $w_t \equiv (w_{11t}, \cdots, w_{1Jt}, \cdots, w_{IJt},)$ has IJ components. Then let $\mathcal{W}$ denote the support of $(X_t^{(1)}, Z_t^{(1)}, w_t)$.

Let $\delta_{ijt}^x = X_{ijt}\alpha$ and $\delta_{ijt}^z = Z_{ijt}\beta$. Define $\delta_{ijt} = (\delta_{ijt}^x, \delta_{ijt}^z)$ and δ_t be the vector of all δ_{ijt}, $(i,j) \in \mathbf{I} \times \mathbf{J}$. Then we can write

$$H(\cdots, X_{ijt}\alpha_{i0}, \cdots, Z_{ijt}\beta_{j0}, \cdots) = H(\cdots, \delta_{ijt}^x + X_{ijt}^{(2)}(\tilde{\alpha}_{i0} - \tilde{\alpha}_i), \cdots, \delta_{ijt}^z + Z_{ijt}^{(2)}(\tilde{\beta}_{0j} - \tilde{\beta}_j), \cdots).$$

Combining this with equation (11) we get that

$$H(\cdots, \delta_{ijt}^x + X_{ijt}^{(2)}(\tilde{\alpha}_{i0} - \tilde{\alpha}_i), \cdots, \delta_{ijt}^z + Z_{ijt}^{(2)}(\tilde{\beta}_{j0} - \tilde{\beta}_j), \cdots) = \phi(\delta_t), \tag{12}$$

where $\phi(\delta_t) = E[Y_t|\delta_t]$. Note that for any δ_t and w_t we can compute x_t and z_t given α and β. So we can view the left hand side of the equation as a function of δ_t and w_t only.

First let all the regressors be continuous, we can choose a point $(\bar{X}^{(1)}, \bar{Z}^{(1)}, \bar{w})$ in the interior of $\mathcal{W}$. Fix $\bar{\delta}$ to be δ_t evaluated at $(\bar{X}^{(1)}, \bar{Z}^{(1)}, \bar{w})$. Since all of $X^{(1)}, Z^{(1)}, w$ consist of continuous variables, it is possible to have a neighborhood $N(\bar{w})$ in the support of w such that equation (12) holds for all $w \in N(\bar{w})$ conditional on $\bar{\delta}$. That is we have

$$H(\cdots, \bar{\delta}_{ijt}^x + X_{ijt}^{(2)}(\tilde{\alpha}_{i0} - \tilde{\alpha}_i), \cdots, \bar{\delta}_{ijt}^z + Z_{ijt}^{(2)}(\tilde{\beta}_{j0} - \tilde{\beta}_j), \cdots) = \phi(\bar{\delta}_t), \quad \forall\, w \in N(\bar{w}).$$

Differentiating the above equation with respect to w and evaluating at $\bar{w}$ we get, for each $q \in \{\mathbf{a}, \mathbf{b}\}$, and $(i,j) \in \mathbf{M}$ we have that

$$\sum_{r=1}^{J} H_{ij,r}^q(\bar{X}\alpha_0, \bar{Z}\beta_0)(\tilde{\alpha}_{10} - \tilde{\alpha}_1) + \cdots + \sum_{r=(I-1)J+1}^{IJ} H_{ij,r}^q(\bar{X}\alpha_0, \bar{Z}\beta_0)(\tilde{\alpha}_{I0} - \tilde{\alpha}_I) +$$

$$+ \sum_{r=IJ+1}^{IJ+I} H_{ij,r}^q(\bar{X}\alpha_0, \bar{Z}\beta_0)(\tilde{\beta}_{10} - \tilde{\beta}_1) + \cdots + \sum_{r=2IJ-I+1}^{2IJ} H_{ij,r}^q(\bar{X}\alpha_0, \bar{Z}\beta_0)(\tilde{\beta}_{J0} - \tilde{\beta}_J) = 0.$$

Since $(\bar{X}^{(1)}, \bar{Z}^{(1)}, \bar{w})$ is an arbitrary point in the interior of $\mathcal{W}$ and measure of boundary of $\mathcal{W}$ is zero, the above equation holds with probability one on $\mathcal{W}$. Then by Assumption 4.3(b) we must have $\alpha_0 = \alpha$ and $\beta_0 = \beta$.

Now suppose we allow w_{ijt} to contain discrete regressors. Then by above arguments the coefficients corresponding to continuous regressors are identified as before. Note that equation (12) holds regardless of whether regressors in w_{ijt} are continuous or discrete. Thus we are left with,

$$
\begin{aligned}
&H(\cdots, X_{ijt}\alpha_i + \rho^x_{ijt}(\tilde{\tilde{\alpha}}_{i0} - \tilde{\tilde{\alpha}}_i), \cdots, Z_{ijt}\beta_j + \rho^z_{ijt}(\tilde{\tilde{\beta}}_{j0} - \tilde{\tilde{\beta}}_j), \cdots) \\
&= H(\cdots, X_{ijt}\alpha_i + \bar{\rho}^x_{ijt}(\tilde{\tilde{\alpha}}_{i0} - \tilde{\tilde{\alpha}}_i), \cdots, Z_{ijt}\beta_j + \bar{\rho}^z_{ijt}(\tilde{\tilde{\beta}}_{j0} - \tilde{\tilde{\beta}}_j), \cdots) \\
&= \phi(\delta_t).
\end{aligned}
$$

Then the result follows from Assumption 4.4 ■

Proof of Proposition 4.6:

The parameters α and β are identified. First consider the case where $\mathcal{X}_1 = \mathbb{R}$ and $\mathcal{Z}_1 = \mathbb{R}$. Then $X_{ijt}\alpha_i = X^{(1)}_{ijt} + X^{(2)}_{ijt}\tilde{\alpha}_i$ and $Z_{ijt}\alpha_i = Z^{(1)}_{ijt} + Z^{(2)}_{ijt}\tilde{\beta}_j$. Consider any point

$$r \equiv (r^x_{11}, r^x_{12}, \cdots, r^x_{IJ}, r^z_{11}, r^z_{12} \cdots, r^z_{IJ}) \in \mathbb{R}^{2IJ}.$$

Fix value of $X^{(2)}_{ijt}\tilde{\alpha}_i = \phi_{ij}$ and $Z^{(2)}_{ijt}\tilde{\beta}_j = \psi_{ij}$. Then there will always exist $X^{(1)}_{ijt} = r^x_{ij} - \phi_{ij}$ and $Z^{(1)}_{ijt} = r^z_{ij} - \psi_{ij}$, such that every component H^q_{ij} in H is identified at point r. Since $r \in \mathbb{R}^{2IJ}$ is arbitrary, H is identified.

Alternatively, in the second part of the proposition, $\mathcal{X}$ and $\mathcal{Z}$ are assumed to be compact subset of $\mathbb{R}^{d_x}$. Then $\mathcal{X}\alpha_i$ is a continuous mapping from $\mathcal{X}$ to $\mathbb{R}$, for all $i \in \mathbf{I}$, and $\mathcal{Z}\beta_j$ is a continuous mapping from $\mathcal{Z}$ to $\mathbb{R}$, for all $j \in \mathbf{J}$. Then also $\mathcal{X}\alpha_i$ and $\mathcal{Z}\beta_j$ are compact subsets of $\mathbb{R}$. Therefore, H is identified on the compact subset $(\mathcal{X}\alpha_1)^J \times \cdots \times (\mathcal{X}\alpha_I)^J \times (\mathcal{Z}\beta_1)^I \times \cdots \times (\mathcal{Z}\beta_J)^I$ of $\mathbb{R}^{2IJ}$. By Assumption 4.2 this compact set has positive measure since at least one of the covariates of both $\mathcal{X}$ and $\mathcal{Z}$ are continuously distributed on its support. Therefore, $\mathcal{H}_I \subsetneq \mathcal{H}$. ■