Unobserved Heterogeneity in Matching Games*

Jeremy T. Fox
University of Michigan and NBER

Chenyu Yang
University of Michigan

April 2011

Abstract

Agents in two-sided matching games are heterogeneous in characteristics that are both observable and unobservable in typical data on matching markets. We investigate the identification of distributions of unobserved agent characteristics, unobserved agent preferences and related complementarity functions using data on who matches with whom. We impose the equilibrium concept of pairwise stability and show that the distribution of unobservables can be recovered. We explore both one-to-one and many-to-many matching games.

*Our email addresses are jeremyfox@gmail.com and chny.yang@gmail.com.

1 Introduction

Matching games model the sorting of agents to each other. Men sort to women in marriage based on characteristics such as income, schooling, personality and physical appearance, with more desirable men typically matching to more desirable women. Upstream firms sort to downstream firms based on the product qualities and capacities of each of the firms. Matching games are distinct from demand models because agents on both sides of the market can be rivals to match with the most attractive partners on the other side of the market. Men compete with other men to marry the most desirable women, and women compete for men. Electronics manufacturers compete for the scarce supplies of newly released components.

There has been recent interest in the structural estimation of two-sided matching games (Dagsvik, 2000; Boyd et al., 2003; Choo and Siow, 2006; Sørensen, 2007; Fox, 2010a; Gordon and Knight, 2009; Chen, 2009; Park, 2008; Yang et al., 2009; Levine, 2009; Echenique et al., 2010), among others. The papers we cite are unified in estimating some aspect of the preferences of agents in a matching game from data on who matched with whom as well as the observed characteristics of agents or of matches. The sorting pattern in the data combined with assumptions about equilibrium inform the researcher about the structural primitives in the market, namely some function that transforms an agent's own characteristics and his potential partner's characteristics into some notion of utility or output. These papers are related to but not special cases of papers estimating noncooperative (Nash) games, like the entry literature in industrial organization (Bresnahan and Reiss, 1991; Berry, 1992). Matching games typically use the cooperative equilibrium concept of pairwise stability.

The empirical literature cited previously has naturally focused on how the preferences of both sides of the market (say men and women) are functions of the characteristics of men and women *observed in the data*. For example, Choo and Siow (2006) study the marriage market in the United States and estimate how the preferences of men for women vary by the age of the man and woman, and likewise for the preferences of women. Fox (2010a) studies matching between automotive assemblers (downstream firms) and car parts suppliers (upstream firms) and asks how observed specialization measures in the portfolios of car parts sourced or supplied contribute to firm payoffs.

Most of the above papers have a relatively limited set of agent characteristics. In Choo and Siow, personality and physical attractiveness are not measured, even if those characteristics are important in real-life marriages. In Fox, each firm's product quality is not directly measured and is only indirectly inferred. If matching based on observed characteristics is found to be important, it is a reasonable conjecture that matching based on *unobserved* characteristics is also important. Ackerberg and Botticini (2002) do not formally estimate a matching game, but they provide empirical evidence that farmers and landlords sort on unobservables such as risk aversion and monitoring ability.

Motivated by the empirical applications cited above, there is a pressing need to estimate matching games with richly specified distributions of unobserved agent heterogeneity. Consider Sørensen (2007), who studies the matching of venture capitalists to entrepreneurs. The rhetoric in the paper is that better venture capitalists may sort to better entrepreneurs, so that the success of a venture capitalist may be from getting access to better investment opportunities and not from adding value to a startup firm. However, the error terms in the estimated matching game do not capture venture capitalist ability or entrepreneur quality. The errors are modeled at the match and not the agent level, and are assumed to be distributed independently across potential matches. A model of unobserved entrepreneur quality, say, would require that all matches involving the same entrepreneur (with different venture capitalists) have positively correlated unobservables. The same assumption of i.i.d. errors at the match level permeates the above literature.

One should not overly criticize a new literature for focusing on sorting on observables instead of unobservables. Indeed, it is not even clear that data on only who matches with whom will be enough to identify distributions of unobserved agent characteristics. This paper investigates whether data on the sorting patterns between agents in matching markets can tell us about the distributions of unobserved agent characteristics relevant for sorting. Thus, this paper is on the nonparametric identification of distributions of unobserved agent heterogeneity in two-sided matching games. If the payoff to a marriage is

$$\text{schooling}^{\text{man}} \cdot \text{schooling}^{\text{woman}} + \text{looks}^{\text{man}}\text{looks}^{\text{woman}}$$

and schooling is recorded in the data while looks is not, this paper in part seeks to identify the joint distribution of male and female looks, conditional on the values of male and female schooling, across matching markets (say cities), up to location and scale normalizations as looks is unmeasured. With the distribution of looks, the researcher can make predictions about the sorting based on looks in the data and construct accurate counterfactual predictions about market equilibria if the distributions of schooling and looks shift (say the returns to schooling increase because of technological change).

A related limitation of the existing literature is the focus on preferences that are identical up to observed agent characteristics. In Echenique, Lee and Shum (2010), the authors study aggregate matchings, where all agents with the same observed type (say men with college degrees) are assumed to have the same preferences. Other studies allow preferences to vary with only i.i.d. errors at the match level, reflecting some idiosyncratic attraction between two people but not always idiosyncratic preferences over characteristics such as schooling. No previous paper allows agents to have flexible, unobservably heterogeneous preferences over the observed characteristics of partners. For example, one man with a college degree may value the schooling of his future spouse while an observably identical man may value his spouse's religion more. Indeed, Hitsch, Hortaçsu and Ariely (2010) use special data from online dating on acceptable and rejected profiles on potential dates to document the extremely heterogeneous preferences of men for women and women for men. This paper allows for this empirically relevant heterogeneity in partner preferences but uses data on only who matches with whom, not data from an online site on rejected profiles.

This paper is motivated primarily by applications in industrial organization, such as the matching of upstream firms to downstream firms. Downstream firms typically pay upstream firms money and it is reasonable to model these transfers as arising in a matching equilibrium, so transferable utility is the reasonable benchmark model. For this reason, we work with transferable utility matching games (Koopmans and Beckmann, 1957; Becker, 1973; Shapley and Shubik, 1972; Ostrovsky, 2008; Hatfield et al., 2011). Much but not all of the literature on marriage also uses transferable utility (Becker, 1973). This paper will not work with data on transfers even though the matching game uses transfers as part of the equilibrium. The reason is data availability: in the marriage study of Choo and Siow (2006), the venture capitalist study of Sørensen (2007) (which did not use transferable utility), and the car parts supplier study of Fox (2010a), the data do not contain the equilibrium transfers between matched parties. For example, in Fox, the monies paid by automobile assemblers to car parts suppliers are private contractual details, even if the suppliers of car parts can be publicly seen by opening the hood of one's car.

The ability to identify a distribution of unobserved agent characteristics is mathematically related to identifying a random effect, random coefficient or mixture distribution (all synonyms here) in other types of economic models. Here, the restrictions from the equilibrium concept in a matching game (pairwise stability) will be used, with the goal being to learn about a distribution unobserved of agent characteristics.

As the key object of empirical interest is a distribution of unobservables, this paper uses as a building block techniques for showing the identification of such distributions of heterogeneity in simpler, single-agent choice models. In particular, our identification proofs work by verifying a sufficient (but not necessary) condition for identification known as separability, which is introduced in Fox and Gandhi (2010). Separability is a property of the economic model and the variation in the independent variables and dependent variables, in matching variation in the observed agent characteristics and the observed matches. In Fox and Gandhi, the main technical assumption motivating the result that separability of an economic model is sufficient for the identification of the distribution of unobservables is that the true distribution must take on at most a finite number of support points on a potentially infinite-dimensional master space of unobservables. The analysis does not assume that the number of support points or their locations are known to the researcher; these objects are learned in identification. Thus, the distribution to be identified lies in an infinite-dimensional space. The results in Fox and Gandhi and in many related papers on identification of qualitative outcome models rely on regressors with large support to ensure that observables will make all unobservable types take all actions. This is not identification at infinity, merely a realization that large values of observables are necessary to get agents with large values of unobservables to shift actions. We use at least one regressor with large support in all of our results for the same reason: a match is a qualitative outcome and large regressor values may be necessary for matches with otherwise low output to occur.

Unlike much of the literature on random effects, we explore the identification of conditional distributions, where the distribution of unobservable characteristics and preferences varies with the realized values of observable characteristics. Thus, we identify a distribution of, say, looks conditional on schooling, which is more analogous to a fixed effects problem. We exploit recent results on discrete mixtures to allow the frequencies of unobservables to depend on observables. Therefore, our identification approach does not suffer from omitted variable bias if observables and unobservables are correlated.

The only other paper to formally explore identification in matching games is Fox (2010b), who investigates the nonparametric identification of features of production functions, which take as arguments the observed characteristics of agents. It is unclear what economic primitives one can learn about with data on who matches with whom, because the researcher has data on only equilibrium matches and not on the choice sets of all agents in equilibrium. As the researcher does not have data on choice sets, standard Nash identification arguments based on the individual rationality premise that each agent maximizes profits subject to actions sets are not applicable. In Fox, the focus is not on primitively modeling error distributions. He models the error terms using a maximum score framework, which, as in single agent maximum score, does not seek to identify the distribution of unobservables (Manski, 1975). Thus, this paper's focus on unobservable heterogeneity complements Fox's identification results on observable heterogeneity.

In some sense, this paper is ahead of its literature because no papers have parametrically estimated distributions of agent-level unobservables. Despite its nonparametric approach, this paper seeks to blaze new ground and discuss a new topic for economic investigation, rather than to simply loosen parametric restrictions in an existing empirical literature.

We first specify the data generating process and review how verifying separability of matching model is a sufficient condition for establishing identification of the distribution of unobservables. Our unobservables nest both unobserved characteristics and unobserved preferences. We then explore identification results in a simple setting, one-to-one matching games. This is done for expositional simplicity. We distinguish two sets of results on one-to-one matching: identification results using in part data on agents who remain unmatched and identification results where all agents are matched. This corresponds to different sorts of datasets. Later we extend our results to many-to-many matching. We do not assume that the payoffs to agents on one side of the market are additively separable across multiple matches.

2 Matching Games and Identification

We first introduce the specification of preferences in a matching game for two-sided, one-to-one matching with transferable utility (Becker, 1973; Roth and Sotomayor, 1990, Chapter 8). We generalize this to many-to-many matching in Section 4. Because of the later generalization to many-to-many matching, we work with the example of upstream and downstream firms, instead of men and women and marriage. We second define the equilibrium concept, pairwise stability, and then discuss the data generating process. We also define identification and review separability as a sufficient condition for proving identification. We finally discuss two technical concepts: analytic functions as well as minimal and maximal vectors.

2.1 Preferences for Matches

For each upstream firm u, there is a vector of observed (in the data) characteristics x_u^{up}. Likewise, for each downstream firm d there is a vector of observed characteristics x_d^{down}. In addition, sometimes we will refer to characteristics specific to a **physical match** $\langle u, d\rangle$, or $x_{\langle u,d\rangle}$. If u and d match as part of a matching outcome, they form a **full match** $\left\langle u, d, t_{\langle u,d\rangle}\right\rangle$, where $t_{\langle u,d\rangle}$ is the real-valued and possibly negative transfer that d pays u in the outcome. The total **profit** of upstream firm m for this full match $\left\langle m, w, t_{\langle m,w\rangle}\right\rangle$ is, for one case we will study,

$$v_{\langle u,d\rangle}^{\text{up}}\left(x_u^{\text{up}}, x_d^{\text{down}}, x_{\langle u,d\rangle}\right) + \gamma^{\text{up}} z_{\langle u,d\rangle} + t_{\langle u,d\rangle}, \tag{1}$$

where $v_{\langle u,d\rangle}^{\text{up}}$ is a random (heterogeneous) function specific to upstream firm u and its physical match $\langle u, d\rangle$ giving the part of u's payoff from its characteristics x_u^{up}, d's characteristics x_d^{down}, and the characteristics specific to their match, $x_{\langle u,d\rangle}$, as well as the scalar $z_{\langle u,d\rangle}$, which has a scalar homogeneous coefficient γ^{up}. In the literal context of an upstream firm selling components to a downstream firm, $v_{\langle u,d\rangle}^{\text{up}}\left(x_u^{\text{up}}, x_d^{\text{down}}, x_{\langle u,d\rangle}\right) + \gamma_{\langle u,d\rangle}^{\text{up}} z_{\langle u,d\rangle}$ might represent the negative manufacturing costs and $t_{\langle u,d\rangle}$ would represent the revenue from sales. The additively separable way transfers enter total profits makes this a transferable utility matching game. The additively separable way $z_{\langle u,d\rangle}$ enters profits is required because of the qualitative nature of the dependent variable. It will be discussed below.

In what follows we will drop the superscripts in terms such as x_u^{up} and x_d^{down}, writing instead x_u and x_d, where the subscripts convey the meaning of upstream firm and downstream firm. Downstream firm d has a total profit of $v_{\langle u,d\rangle}^{\text{down}}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + \gamma^{\text{down}} z_{\langle u,d\rangle} - t_{\langle u,d\rangle}$ for the full match $\left\langle u, d, t_{\langle u,d\rangle}\right\rangle$. We normalize the non-$t$ payoff of being unmatched to be 0 for all upstream firms and all downstream firms. A physical match when upstream firm u is unmatched is $\langle u, 0\rangle$ and a downstream firm being unmatched is $\langle 0, d\rangle$. An unmatched full match is $\langle u, 0, 0\rangle$ or $\langle 0, d, 0\rangle$.

The rich unobserved heterogeneity this paper seeks to identify is wrapped up in the heterogeneous functions $v_{\langle u,d\rangle}^{\text{up}}$. Indeed, the $v_{\langle u,d\rangle}^{\text{up}}$ are random functions: they vary across upstream firms and across matches involving the same upstream firm and different downstream firms.

In one-to-one matching, it is not possible to separably identify the payoffs of upstream firms and downstream firms. References such as Roth and Sotomayor (1990, Chapter 8) show that the set of matches that occur is a function of the **production** from a match, or

$$v^{\text{up}}_{\langle u,d\rangle}\left(x^{\text{up}}_u, x^{\text{down}}_d, x_{\langle u,d\rangle}\right) + v^{\text{down}}_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + \left(\gamma^{\text{up}} + \gamma^{\text{down}}\right) z_{\langle u,d\rangle}.$$

As the production of matches in monetary units cannot be learned without using data on the equilibrium transfers, which we do not use, we will need location and scale normalizations for identification. When data on unmatched agents are used, the production of a singleton match will be normalized to 0 (there are no $z_{\langle u,d\rangle}$'s for singleton matches). This will serve as the location normalization. For the scale normalization, we set $\left(\gamma^{\text{up}} + \gamma^{\text{down}}\right) = 1$ for the homogeneous coefficients on $z_{\langle u,d\rangle}$. We could also set $\left(\gamma^{\text{up}} + \gamma^{\text{down}}\right) = -1$ or actually even learn the sign of $\left(\gamma^{\text{up}} + \gamma^{\text{down}}\right)$ in identification.[1] Therefore, the production of a match will be in units of $z_{\langle u,d\rangle}$.

In new notation, the production function for match $\langle u, d\rangle$ is

$$f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + z_{\langle u,d\rangle}, \tag{2}$$

where

$$f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) = v^{\text{up}}_{\langle u,d\rangle}\left(x^{\text{up}}_u, x^{\text{down}}_d, x_{\langle u,d\rangle}\right) + v^{\text{down}}_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right)$$

is the **non-z production function**. For one-to-one matching, we will seek to learn the joint distribution of the match-specific production functions $f_{\langle u,d\rangle}$ across all matches $\langle u, d\rangle$.

The general notation $f_{\langle u,d\rangle}$ nests many important examples. One special case of (2) is

$$f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + z_{\langle u,d\rangle} = \left(x_u \cdot x_d\right)' \beta_{\langle u,d\rangle,1} + x'_{\langle u,d\rangle}\beta_{\langle u,d\rangle,2} + e_{\langle u,d\rangle} + z_{\langle u,d\rangle}, \tag{3}$$

where $x_u \cdot x_d$ includes all non-interacted terms and all forms all pairwise interactions between one element of the upstream firm vector x_u and one element of the downstream firm vector x_d, $\beta_{\langle u,d\rangle,1}$ is a vector of random coefficients specific to the match $\langle u, d\rangle$ on these interactions, $\beta_{\langle u,d\rangle,2}$ is a vector of random coefficients for the characteristics $x_{\langle u,d\rangle}$ that vary at the level of the physical match $\langle u, d\rangle$, and $e_{\langle u,d\rangle}$ is an additive unobservable in the payoff of match $\langle u, d\rangle$. The heterogeneous values $\left(\beta_{\langle u,d\rangle,1}, \beta_{\langle u,d\rangle,2}\right)$ give the marginal contribution of each regressor or regressor combination to match $\langle u, d\rangle$'s utility and are the analogs to the random coefficients in discrete choice estimation, like the random coefficient logit (Boyd and Mellman, 1980; Cardell and Dunbar, 1980). We do not assume that $e_{\langle u,d\rangle}$ has mean 0, so $e_{\langle u,d\rangle}$ plays the role of an additive intercept specific to match $\langle u, d\rangle$.

A particular case of the special case (3) is when

$$\left(\beta_{\langle u,d\rangle,1}, \beta_{\langle u,d\rangle,2}\right) = \left(\beta_{u,1}, \beta_{u,2}\right) + \left(\beta_{d,1}, \beta_{d,2}\right),$$

[1] It is trivial to see if raising $z_{\langle u,d\rangle}$ for some match increases the probability of observing assignments containing that match.

or the random coefficients in the production function for match $\langle u, d\rangle$ are decomposed into the *random preferences* in upstream firm u's profits and the random preferences in downstream firm d's profits. As we do not restrict the functions $f_{\langle u,d\rangle}$ to be independently distributed across matches, we allow the important case where upstream firms have heterogeneous preferences for downstream firm and match characteristics and downstream firms have heterogeneous preferences for upstream firm and match characteristics. Turning to marriage and dating for a minute, heterogeneous preferences fit the empirical findings Hitsch et al. (2010), where agents with the same characteristics have been found to have quite different preferences.

Also in (3), we can decompose $e_{\langle u,d\rangle} = e_u \cdot e_d$, where e_u is an *unobserved characteristic* of upstream firm u such as its product quality and e_d is such an unobserved characteristic for downstream firm d. Thus, a special case of our analysis is the unobserved characteristics such as firm quality missing in Fox (2010a) and Sørensen (2007).

Finally, the special case of (3), itself a special case, that is the most obvious extension of the past literature along the lines of the discussion about adding unobserved agent characteristics but not unobserved preferences in the introduction is

$$f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + z_{\langle u,d\rangle} = \left(x_u \cdot x_d\right)' \beta_1 + x'_{\langle u,d\rangle}\beta_2 + e_u \cdot e_d + z_{\langle u,d\rangle}, \tag{4}$$

where now the marginal values (β_1, β_2) of the observed characteristics are homogeneous, and where, as before, e_u is an unobserved upstream firm characteristic and e_d is an unobserved downstream firm characteristic.

2.2 Equilibrium: Pairwise Stability

An **outcome** O to this type of transferable utility matching game is a set of full matches $\left\langle u, d, t_{\langle u,d\rangle}\right\rangle$ such that each firm is matched at most once. The most common equilibrium concept in a matching game is **pairwise stability** (Roth and Sotomayor, 1990, Chapter 8). Loosely speaking, a pairwise stable outcome is one where no pair of firms that are not matched to each other at O prefers to deviate and become matched to each other, no individual firm that is matched wants to become unmatched, and no two unmatched firms want to become matched. Because of the equilibrium transfers, pairwise stability can also be expressed in terms of price taking behavior. In part, an upstream firm u not matched to downstream firm d will not want to match with d and receive the transfer that equalizes d's total profit from u and the profit d gets from d's full match in O. The set of physical matches corresponding to an outcome O will be called an **assignment** $A(O)$ or just A. The assignment portion of a pairwise stable outcome will be called, informally, a **pairwise stable assignment**. An assignment must be feasible: each firm must be matched at most once. For an assignment A, we denote the downstream firm matched to upstream firm u as the value of the mapping $A(u)$.

Any outcome to a one-to-one matching game that is pairwise stable is also **fully stable** (Roth and Sotomayor, 1990, Chapter 8). A fully stable outcome is robust to deviation by any coalition of agents. One such coalition is the coalition of all agents, so a pairwise stable assignment $A(O)$ must maximize the total production in the economy

$$\sum_{\langle u,d\rangle \in A(O)} \left\{ f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) + z_{\langle u,d\rangle} \right\}. \tag{5}$$

While transfers enter an outcome O, they do not enter the production of matches. Therefore, in one-to-one matching one can compute a pairwise stable assignment by maximizing (5) over all assignments A. In other words, we can solve a social planner's multinomial choice problem over assignments. In principle there can be multiple assignments A that maximize (5). The observed regressors x_u, x_d, $x_{\langle u,d\rangle}$ and $z_{\langle u,d\rangle}$ will be assumed to have continuous and product supports. Discrete regressors are mentioned below. If the observed regressors enter nontrivially into production, ties will occur with probability 0 and there will be a unique pairwise stable assignment with probability 1.

2.3 Data Generating Process and Identification

Proving identification shows that one can learn about the objects of interest. Here we will use variation in observed agent characteristics across markets. Indeed, we will observe data on a continuum of matching markets. Consider a matching market. Let $\mathcal{U}$ be the set of upstream firms as well as the number of such firms. Likewise, let $\mathcal{D}$ be the number and set of downstream firms. Let X be the non-z set of observed characteristics of agents and matches

$$X = \left((x_u)_{u\in\mathcal{U}}, (x_d)_{d\in\mathcal{D}}, \left(x_{\langle u,d\rangle}\right)_{u\in\mathcal{U},d\in\mathcal{D}}, \mathcal{U}, \mathcal{D} \right)$$

and let $Z = \left(Z_{\langle u,d\rangle}\right)_{u\in\mathcal{U},d\in\mathcal{D}}$. We have included $\mathcal{U}$ and $\mathcal{D}$ in X to conserve notation. Let θ comprise all unobservables in a market,

$$\theta = \left(\left(f_{\langle u,d\rangle}\right)_{u\in\mathcal{U},d\in\mathcal{D}} \right). \tag{6}$$

The so-called market-level type θ is a realization of $\mathcal{U}\cdot\mathcal{D}$ functions $f_{\langle u,d\rangle}$ of three vectors, x_m, x_w and $x_{\langle m,w\rangle}$. If instead the specification is (3), the definition of θ is

$$\theta = \left(\left(\beta_{\langle u,d\rangle,1}, \beta_{\langle u,d\rangle,2}, e_{\langle u,d\rangle}\right)_{u\in\mathcal{U},d\in\mathcal{D}} \right). \tag{7}$$

Now θ is comprised of $2\cdot\mathcal{U}\cdot\mathcal{D}$ vectors of real numbers such as $\beta_{\langle u,d\rangle,1}$ and $\mathcal{U}\cdot\mathcal{D}$ scalars $e_{\langle u,d\rangle}$. This type θ lies in real space and the more general (6) lies in a function space. We will consider other examples of market-level types, corresponding to the types of objects that can be identified with different types of matching data.

The researcher is assumed to have data on (A, X, Z) from an independent and identically distributed set of markets. In other words, in each matching market we observe who matches with whom A and the characteristics X and Z of the agents and the realized and potential matches. The probability of assignment A occurring given the observables X and Z is

$$\Pr(A \mid X, Z; G) = \int_\theta 1\left[A \,\text{pairwise stable assignment} \mid X, Z, \theta\right] dG(\theta \mid X, Z), \tag{8}$$

where $1\left[A \,\text{pairwise stable assignment} \mid X, Z, \theta\right]$ is equal to 1 when A is a pairwise stable assignment for the market (X, Z, θ) and $G(\theta \mid X, Z)$ is the distribution of market-level unobservables θ conditional on market-level observables (X, Z). We do not impose the assumption that (X, Z) and θ are independently distributed. We allow the distribution of a firm's unobservables to be a function of its observable characteristics.

The goal of this paper is to learn the distribution $G(\theta \mid X, Z)$. Knowledge of G tells us the conditional distribution of unobserved agent characteristics and unobserved agent preferences in a matching game. Some of these agent characteristics and characteristics may not be additively separable terms; indeed they may be functions. The conditional probability $\Pr(A \mid X, Z)$ is identified from data on (A, X, Z), so we can informally see whether the model's predictions for a given G match these data. The distribution G is said to be **identified** whenever, for two $G^1 \neq G^2$, $\Pr\left(A \mid X, Z; G^1\right) \neq \Pr\left(A \mid X, Z; G^2\right)$ for some tuple (A, X, Z). In other words, G^1 and G^2 give a different implication for the conditional probability of at least one assignment A given the observable agent and match characteristics in X and Z. By continuity of the model and the to-be-assumed continuity of $G(\theta \mid X, Z)$ in (X, Z), the existence of one such pair (A, X, Z) implies that a continuum of (X, Z), a set with positive measure, satisfies $\Pr\left(A \mid X, Z; G^1\right) \neq \Pr\left(A \mid X, Z; G^2\right)$.

2.4 The Sufficient Condition of Separability

Verifying directly that (8) has a unique $G(\theta \mid X, Z)$ that rationalizes it is difficult for other, simpler examples, such as the single-agent multinomial choice model. Fox and Gandhi (2010) address this concern for examples like the multinomial choice model by developing a sufficient condition for proving that G is identified in economic choice models. Fox and Gandhi do not consider games but the sufficient condition extends to matching games. The sufficient condition is called separability. The definition in matching notation follows.

Definition 2.1. A two-sided, one-to-one matching game satisfies **separability** if, for **any** set T of a finite number of market-level unobservable types θ, we can find an assignment A and observable characteristics (X, Z) such that only one market unobservable θ in T makes A a pairwise stable assignment at (X, Z).

Note also that the set T in the definition of separability is an arbitrary finite set (of vectors of functions) and is not the support of the true $G(\theta \mid X, Z)$. This paper will investigate the identification of the distribution $G(\theta \mid X, Z)$ of unobserved characteristics by verifying separability for several versions of matching games. To verify separability, one starts with an arbitrary, finite set T of market-level unobservables θ's and then searches for one particular tuple (A, X, Z) where we can isolate one particular $\theta \in T$ and show it alone will make A a pairwise stable assignment with observables (X, Z). Note that this verification of separability is non-trivial in matching games, as the proofs will show. Still, the tool of separability puts the focus on aspects of identification that are specific to matching games rather than on general issues about solving integral equations (8).

Separability leads to identification when the distribution $G(\theta \mid X, Z)$ of θ has finite support independent of (X, Z). In full generality, each market-level type θ lies in an infinite-dimensional functional space Θ. The assumption of finite support means that at most a finite number of vectors of functions takes on positive support for each possible G.

Assumption 2.2.

1. *The distribution G lies in a class $\mathcal{G}$ of distributions on the finite or infinite-dimensional space Θ, where each element of $\mathcal{G}$ has at most a finite number of support points.*

2. *Further, each weight $p_n(X, Z)$ on type n in the distribution G is an analytic function of (X, Z).*

Analytic functions are discussed below. With this assumption, the G corresponding to (6) can be written as

$$G = \left(N, (\theta_n)_{n \in N}, (p_n(X, Z))_{n \in N}\right) = \left(N, \left(\left(f_{\langle u,d \rangle, n}\right)_{u \in \mathcal{U}, d \in \mathcal{D}}\right)_{n \in N}, (p_n(X, Z))_{n \in N}\right), \tag{9}$$

where N is the number of support points, θ_n is the nth support point in function space, and again $p_n(X, Z)$ is the weight on type θ_n. Weights depend on X and Z because we do not assume independence of X and Z from θ. If $\mathcal{U}$ and $\mathcal{D}$ are fixed across markets, we identify $\mathcal{U} \cdot \mathcal{D} \cdot N$ non-z production functions $f_{\langle u,d \rangle}$ and N weighting functions $p_n(X, Z)$, where N is also learned in identification. In other notation, we learn the number of support points N, their locations θ_n and the weights p_n in identification. $\mathcal{G}$ is an infinite-dimensional class of distribution functions in part because each G takes on support on a function space Θ and the weights $p_n(X, Z)$ are also functions of (X, Z).

Proposition 2.3. *Let Assumption 2.2 hold. Then the distribution of unobserved heterogeneity G is identified within the class $\mathcal{G}$ if the model is separable.*

The theorem, without conditioning G on (X, Z), is also found in Fox and Gandhi. A proof is included in the Appendix.

Under (9), the data generating process (8) becomes

$$\Pr(A \mid X, Z; G) = \sum_{n=1}^{N} 1\left[A \, \text{pairwise stable assignment} \mid X, Z, \theta_n\right] p_n(X, Z). \tag{10}$$

Even if we assumed θ is distributed independently of (X, Z), we cannot simply run a linear regression to estimate the now constant p_n's as the number N of support points (market-level vectors of functions) and their locations θ_n also need to be learned.

2.5 Support Conditions and Analytic Functions

Throughout the paper, the definition of X and Z may change as the available matching data change. Regardless of the exact definitions, we will always maintain the following support condition.

Assumption 2.4. *The support of Z is the product space $\mathbb{R} \times \ldots \times \mathbb{R}$ and the support of X is an arbitrarily small open set in $\mathbb{R} \times \ldots \times \mathbb{R}$.*

We need large support on the elements of Z in order to induce agents to undertake certain matches or assignments. This is similar to the need for large support in models of multinomial choice, where large support in one regressor per choice (say charging a high price for a product) is needed to induce agents with a high value for that product to purchase another product instead, so we can identify their preference (Ichimura and Thompson, 1998; Berry and Haile, 2010b; Fox and Gandhi, 2010).

We do not need large support on X because we will restrict the production functions to be analytic functions.

Assumption 2.5. *All heterogeneous functions in each θ, such as the non-z production functions, are analytic functions.*

More properly, they are multivariate real analytic functions.[2] Two analytic functions take on different values on an arbitrarily small open set, which is why we do not need large support on the elements of X that enter each $f_{\langle u,d\rangle}$ in θ. More generally, there exists a point X in any open set where any finite number of analytic functions will all take on different values (Fox and Gandhi, 2010). This property is also used for $P(Z, X)$. Instead of explicitly assuming the functions are analytic, one could instead assume that "for any finite set T of market level observables θ and for $\theta_1, \theta_2 \in T$, there exists a point X in any open set where $f^1_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right) \neq f^2_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right)$ for one match production function $\langle u, d\rangle$, and $f^1_{\langle u,d\rangle}$ is the production function for realization θ_1, as is $f^2_{\langle u,d\rangle}$ for θ_2." Analytic functions would be a sufficient condition for this "no-ties" property.

Distributions of analytic functions, as in (2), provide important generalities over distributions of finite-dimensional vectors, as in (3). Over a space of functions, a researcher does not have to take a stand on functional form, like whether a regressor enters linearly as in x, as an interaction as in x_1x_2, as a power such as x^2, as a real power $x^{2.23}$, as a logarithm $\log(x)$, or in more unusual ways such as $\sin\left(\log\left(22x_1 + x_2^{2.33}\right)\right)$. Restricting attention to one of either discrete distributions over functions or, if the results were extended, continuous distributions over a vector of finite-dimensional, real

[2] An analytic function is always locally equal to a power series.

parameters may be important as Ackerberg, Hahn and Ridder (2010) present a related production function example where continuous distributions of functions are not identified but where Fox and Gandhi (2010) would be able to show countable distributions of functions are identified.

We do not identify heterogeneous functions of discrete covariates, as in the literature on multinomial choice, which cannot in fully generality identify distributions of heterogeneous coefficients or heterogeneous functions on discrete covariates. We can condition all arguments on discrete covariates, so that we can identify a distribution $G(\theta \mid X, Z, D)$, where D is the vector of observable discrete characteristics for all agents and matches and the entire set of vectors of functions θ that take on positive support can shift with D. We suppress discrete characteristics for convenience in what follows. Unlike the work of Berry and Haile (2010b,a) on identification in multinomial choice models, we identify distributions over functions of continuous characteristics, not just utility values conditional on continuous characteristics.

2.6 Minimal and Maximal Vectors

In our proofs of separability and hence identification of $G(\theta \mid X, Z)$ for various specifications, we often work with what we call a **minimal** or **maximal** vector. Given a finite set B of vectors b, a minimal vector $b_1 \in B$ is one such that every $b \in B \setminus \{b_1\}$ has at least one element greater than the corresponding element of b_1. In other words, $b_1 \ngeq b$ for all $b \in B \setminus \{b_1\}$ by the usual partial product order on vectors. Given the set B, there may be multiple minimal vectors, but there always exists one by finiteness. Likewise, a maximal vector b_1 is such that $b_1 \nleq b$ for all $b \in B \setminus \{b_1\}$ by the usual product order on vectors.

3 Identifying Heterogeneity for One-to-One Matching

This section examines the identification of distributions of heterogeneity in the relatively simple one-to-one matching games outlined in Section 2. Identification intuitively occurs by observing different markets with different characteristics (X, Z) and seeing how matching assignment probabilities $\Pr(A \mid X, Z)$ vary. In other words, the sorting pattern variability across markets as a function of observable characteristics tells us about the importance of unobservables.

The arguments work by verifying separability. Thus we find some (X, Z) at which one vector of market unobservable functions or random coefficients θ will give a different assignment A than any other of a finite set of alternative θ's. Intuitively, verifying separability is about showing that matching markets with different values of the unobservables have different assignments, meaning different sorting patterns.

We will consider four cases with positive identification results. First, each case is distinguished by whether the elements of Z are the match-specific $z_{\langle u,d \rangle}$ or whether we instead work with firm-specific characteristics z_u and z_d that enter into production multiplicatively, meaning $z_u \cdot z_d$. Second, each case is distinguished by whether the analysis allows firms to be unmatched and uses data on unmatched firms. In marriage, having data on single people is common, but not all empirical work on business to business matching has data on potential entrants, a name for firms unmatched in a pairwise stable outcome. The four cases we study identification for are combinations of the two dimensions of 1) match-specific or firm-specific elements of Z and 2) whether agents can be unmatched or all agents are required to be matched.

3.1 Identification Using Data on Unmatched Agents

We start with an upstream-downstream market where the researcher observes various geographically (or temporally, or industrially) separated matching markets. In this section, in each matching market the researcher observes both the identities of the firms that are unmatched (say potential entrants) and the match partners of firms that are matched.

Let the production to a match be given by (2) and so θ is given by (6). This is the case where each match has a covariate $z_{\langle u,d\rangle}$. Recall that this case allows for the most heterogeneity and nests unobserved agent preferences and unobserved agent characteristics. The proof is simple and identification relies heavily on the decision to be unmatched instead of matched.

Theorem 3.1. *The distribution $G(\theta \mid X, Z)$ is identified in the one-to-one matching game with unmatched agents and match-specific unobservable production functions.*

Proof. We prove separability. We are given an arbitrary T, which is comprised of a finite number of market-level unobservables θ's. For any θ and X, let $w(\theta, X)$ be

$$w(\theta, X) = \left(f_{\langle 1,1\rangle}\left(x_1^{\text{up}}, x_1^{\text{down}}, x_{\langle 1,1\rangle}\right), \ldots, f_{\langle \mathcal{U},\mathcal{D}\rangle}\left(x_{\mathcal{U}}, x_{\mathcal{D}}, x_{\langle \mathcal{U},\mathcal{D}\rangle}\right)\right),$$

the list of non-z production functions in θ, as functions of the elements of X. The vector $w(\theta, X)$ is itself a vector of analytic functions and each θ_1 has at least one function different in $w(\theta_1, X)$ than one function in $w(\theta_2, X)$, by the notion that two types θ_1 and θ_2 are different. By the properties of analytic functions, there exists some X in its support such that the values $w(\theta_1, X) \neq w(\theta_2, X)$, for any $\theta_1 \neq \theta_2$. Then there exists a minimal vector $w(\theta_1, X)$ at that X. For each match, set $z_{\langle u,d\rangle} = -f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right)$ for θ_1. Then for θ_1, the production of each match is 0 and for any other $\theta \in T$ at least one match has positive production. Under θ_1 and this (X, Z), every agent will be single. Under any other $\theta \in T, \theta \neq \theta_1$, at least one non-single match will occur. Thus, the model is separable. □

3.2 Agent-Specific Unobservables and Unmatched Agents

Many datasets may not have a match-specific regressor $z_{\langle u,d\rangle}$ with large support. In this section, we instead assume that regressors with large support varies at the agent level, which may be more common in certain datasets.

Let the production to a match involving upstream firm u and downstream firm d be

$$g_u(x_u) \cdot g_d(x_d) + z_u \cdot z_d, \tag{11}$$

where g_u is a heterogeneous **index function**, for lack of a better label, for u and g_d is such an index function for d. The index functions for u and d combine to make a non-z production function for that match. The idea is that u's index function g_u is an index that captures the total contribution of its observed and unobserved characteristics to production. A special case of an index function is

$$g_u(x_m) = x_u' \beta_u + e_u,$$

where β_u is a heterogeneous vector of random coefficients governing the importance of upstream firm u's non-z observed characteristics in production and e_u is an additive unobserved characteristic for upstream firm u. The key idea is that the non-z observed characteristics of a u are combined into a single index that interacts with the non-z characteristics of downstream firms. The weight β_u on x_u is heterogeneous across firms.

We impose the additional scale normalization that $g_1(x_1^{\text{up}}) \equiv 1$ for all values of x_1^{up}, for upstream firm $u = 1$. The market-level unobservable is therefore

$$\theta = \left((g_u)_{u=2}^{\mathcal{U}}, (g_d)_{d \in \mathcal{D}} \right).$$

Our goal is to identify a distribution $G(\theta \mid X, Z)$ over market-level types comprised of $\mathcal{U} - 1 + \mathcal{D}$ functions. Without the scale normalization $g_1(x_1^{\text{up}}) \equiv 1$, we could multiply all downstream firm g_d's by a constant and divide all g_u's by the same constant, leaving the production (11) of all matches the same. We also continue to maintain the location normalization that the value of being unmatched is 0.

Theorem 3.2. *The distribution $G(\theta)$ is identified in the one-to-one matching game with unmatched agents and agent-specific index functions.*

Proof. We prove separability. We are given an arbitrary T, which is comprised of a finite number of θ's. Let $w(\theta, X)$ be the long vector of production from each match, other than from $z_u \cdot z_d$, meaning

$$w(\theta, X) = (1 \cdot g_1(x_1), \cdots, g_{\mathcal{U}}(x_{\mathcal{U}}) \cdot g_{\mathcal{D}}(x_{\mathcal{D}})).$$

First we prove that if $\theta_1 \neq \theta_2$, then $w(\theta_1, X) \neq w(\theta_2, X)$ as functions of X. Because $\theta_1 \neq \theta_2$, at least one g_u or g_d is different between θ_1 and θ_2. Let g_d differ between θ_1 and θ_2. Then the match $\langle 1, d \rangle$ has a different production function $1 \cdot g_d(x_d)$. Now say the g_d for all $d \in \mathcal{D}$ are the same between θ_1 and θ_2 and at least one g_u is different. Then the match $\langle u, d \rangle$ has a different non-z production function $g_u(x_u) \cdot g_d(x_d)$, as again g_u is different and g_d is the same between θ_1 and θ_2.

As $w(\theta_1, X) \neq w(\theta_2, X)$ as functions of X and as the product of analytic functions are analytic, there exists an X in its support such that $w(\theta_1, X) \neq w(\theta_2, X)$ for all $\theta_1, \theta_2 \in T$ at that X. Then there exists a minimal vector $w(\theta_1, X)$. Let $z_u = -g_u(x_u)$ and $z_d = g_d(x_d)$ for all u and d and the g's being elements of θ_1. Then at θ_1 and this X, the production from all individual matches is 0. So all agents will be unmatched at this θ_1. All other $\theta \in T$ will have some matches with strictly positive production, so that the agents in one or more of those matches will not be unmatched. Thus, only θ_1 gives the assignment of all agents being unmatched and the model is separable. □

3.3 Unmatched Agents: Non-Identification Results

Return to the specification (2), where the covariates in Z are of the match-specific $z_{\langle u,d \rangle}$ nature. The identification results in the previous section rely heavily on the individual rationality decision to be unmatched. As the results and their relatively simple proofs show, data on unmatched agents are a powerful source of identification. However, data on matching markets often lack data on unmatched agents. For example, in the automotive application of Fox (2010a) there are no data on potential but not actual entrants to supplying car parts or assembling cars. One would like to proceed with identification without relying on the individual rationality decision to be unmatched. In what follows, we work with matching markets where agents are unable to be unmatched. We rely only on the sorting pattens of matched agents to each other for identification.

To avoid unmatched agents because of market imbalance, we set the number of upstream firms to be equal to the number of downstream firms, or $\mathcal{U} = \mathcal{D}$. One-to-one matching markets have assignments that solve the social planner's problem (5). Thus, to show that a model is separable and hence $G(\theta \mid X, Z)$ is identified, we need to, in part, show that there exists some X where two θ's have a different sum of non-z production for at least one assignment.

First we argue that the distribution of the types of market-level unobserved characteristics that we have described is not directly identifiable using matching markets without unmatched agents. For an example where we will show non-identification, consider a simple model where production is

$$z_{\langle u,d \rangle} + e_{\langle u,d \rangle},$$

where $e_{\langle u,d \rangle}$ is unobservable, so that the market level unobservable is

$$\theta = \left(e_{\langle u,d \rangle}\right)_{u \in \mathcal{U}, d \in \mathcal{D}}.$$

Dropping X from the model, an important step of the proof in Section 3.1 is to show, for the long vector of non-z production for each match

$$w(\theta) = \left(e_{\langle 1,1\rangle}, \ldots, e_{\langle \mathcal{U},\mathcal{D}\rangle}\right),$$

that $w(\theta_1) \neq w(\theta_2)$ for any $\theta_1 \neq \theta_2$. A notable feature of the examples of non-identification here is that, in the absence of the individual rationality decision to remain single, the intermediate result that at least one match has different production levels for θ_1 and θ_2 is not sufficient to show identification of $G(\theta)$.

Example 3.3. Let the matching market involve two upstream firms and two downstream firms. The vector of the unobservable components of production for each match is therefore

$$w(\theta) = \left(e_{\langle 1,1\rangle}, e_{\langle 1,2\rangle}, e_{\langle 2,1\rangle}, e_{\langle 2,2\rangle}\right).$$

Denote the production unobservables of two different types θ_1, θ_2 as $e^1_{\langle u,d\rangle}$ and $e^2_{\langle u,d\rangle}$. Now suppose

$$\begin{aligned} e^1_{\langle 1,1\rangle} + e^1_{\langle 2,2\rangle} = e^2_{\langle 1,1\rangle} + e^2_{\langle 2,2\rangle} = Y_1 \\ e^1_{\langle 1,2\rangle} + e^1_{\langle 2,1\rangle} = e^2_{\langle 1,2\rangle} + e^2_{\langle 2,1\rangle} = Y_2 \end{aligned}$$

for two scalars Y_1, Y_2. This condition does not preclude $w(\theta_1) \neq w(\theta_2)$. However, given any scalar characteristics $Z = \left(z_{\langle u,d\rangle}\right)_{u\in\mathcal{U}, d\in\mathcal{D}}$, the stable equilibrium assignment will be the same under θ_1 and θ_2. To see this, notice that if

$$z_{\langle 1,1\rangle} + z_{\langle 2,2\rangle} + Y_1 > z_{\langle 1,2\rangle} + z_{\langle 2,1\rangle} + Y_2,$$

the pairwise stable assignment is $\{\langle 1,1\rangle, \langle 2,2\rangle\}$ for both θ_1 and θ_2. When the inequality is reversed, the outcome is $\{\langle 1,2\rangle, \langle 2,1\rangle\}$ for both θ_1 and θ_2. Therefore there does not exist a set of $z_{\langle m,w\rangle}$ that can distinguish θ_1 from θ_2.

Suppose θ_1 and θ_2 satisfy the relationship above, and two distributions $G_a(\theta \mid X, Z)$ and $G_b(\theta \mid X, Z)$ take on different positive probabilities on θ_1 and θ_2, and are 0 everywhere else. Then the resulting conditional probabilities $\Pr(A \mid X, Z; G)$ are the same for all (A, X, Z), for both G_a and G_b. Hence, the model that includes G_a and G_b in the distribution space $\mathcal{G}$ and types θ_1, θ_2 in the space of market-level unobservables Θ is not identified. Also, if the model is not identifiable it cannot be shown to be separable.

We can generalize the results above to the wider class of one-to-one matching markets involving $\mathcal{U}$ upstream firms and $\mathcal{D} = \mathcal{U}$ downstream firms. Because all agents are matched, the production for a matched pair is not restricted to be nonnegative. Instead, we normalize the unobserved component of the production of matches involving an upstream firm and a downstream of equal index ($u = d$)

to be 0: $e_{\langle u,u\rangle} = 0, \forall u = 1, \ldots, \mathcal{U}$. This location normalization does not affect the pairwise stable assignments given θ and Z: we can add or subtract a constant to the production of each match involving upstream firm u and the sum of production of every assignment will change by an equal amount, as u is matched exactly once in each assignment. Therefore, the assignment that maximizes the total production does not change.

Example 3.4. For convenience, we write each θ as a matrix in this simple production function structure. Again, let $\mathcal{D} = \mathcal{U}$ and let rows index upstream firms and columns downstream firms:

$$\theta_1 = \begin{pmatrix} 0 & e_{\langle 1,2\rangle} & \dots & e_{\langle 1,\mathcal{U}\rangle} \\ e_{\langle 2,1\rangle} & 0 & \dots & \epsilon_{\langle 2,\mathcal{U}\rangle} \\ \vdots & \vdots & \ddots & \vdots \\ e_{\langle \mathcal{U},1\rangle} & e_{\langle \mathcal{U},2\rangle} & \dots & 0 \end{pmatrix}, \theta_2 = \begin{pmatrix} 0 & e_{\langle 1,2\rangle} + 1 & \dots & e_{\langle 1,\mathcal{U}\rangle} \\ e_{\langle 2,1\rangle} - 1 & 0 & \dots & \epsilon_{\langle 2,\mathcal{U}\rangle} - 1 \\ \vdots & \vdots & \ddots & \vdots \\ e_{\langle \mathcal{U},1\rangle} & e_{\langle \mathcal{U},2\rangle} + 1 & \dots & 0 \end{pmatrix}.$$

To form θ_2 from θ_1, we add 1 to all matches involving downstream firm 2 and subtract 1 from all matches involving upstream firm 2. Clearly, $w(\theta_1) \neq w(\theta_2)$: some matches have different non-z production levels under θ_1 and θ_2. However, the two types θ_1 and θ_2 have the same sums of production for all assignments. To see this, notice that any assignment involves either the match $\langle 2, 2\rangle$ or two matches of the form $(\langle 2, d\rangle, \langle u, 2\rangle)$. In the first case, $e_{\langle 2,2\rangle} = 0$ in both θ_1 and θ_2; in the second case, the sum of the production levels for any assignment is still the same for θ_1 and θ_2, as $+1 - 1 = 0$. By an argument similar to the previous example, we can show that the distribution over the two types is not identifiable and therefore is not separable.

3.4 Matched Agents: Identification of the Distribution of Complementarity Functions

The examples in the previous section show that it is not always possible to identify the distribution of the most primitive unobserved heterogeneity in matching games. Remaining with the same underlying model, here we focus on identifying the distribution of, for lack of a more precise label, **complementarity functions**. The complementarity function involving upstream firms u_1, u_2 and woman d_1, d_2 is defined to be

$$\begin{aligned} c(u_1, u_2, d_1, d_2, X) = {} & f_{\langle u_1,d_1\rangle}\left(x_{u_1}, x_{d_1}, x_{\langle u_1,d_1\rangle}\right) + f_{\langle u_2,d_2\rangle}\left(x_{u_2}, x_{d_2}, x_{\langle u_2,d_2\rangle}\right) \\ & - \left(f_{\langle u_1,d_2\rangle}\left(x_{u_1}, x_{d_2}, x_{\langle u_1,d_2\rangle}\right) + f_{\langle u_2,d_1\rangle}\left(x_{u_2}, x_{d_1}, x_{\langle u_2,d_1\rangle}\right)\right), \end{aligned} \quad (12)$$

where a non-z production function is in (2). The arguments of a complementarity function are

$$\left(x_{u_1}, x_{d_1}, x_{u_2}, x_{d_2}, x_{\langle u_1,d_1\rangle}, x_{\langle u_1,d_2\rangle}, x_{\langle u_2,d_1\rangle}, x_{\langle u_2,d_2\rangle}\right)$$

and the function varies heterogeneously across pairs of two upstream firms and two downstream firms. Notationally, we instead write the arguments of $c(u_1, u_2, d_1, d_2, X)$ as the agent indices and the master list of non-z observables X in order to reduce the list of arguments. In a one-to-one matching game and given X, the complementarity function c tells us which outcome is more desirable in a two-by-two submarket, ignoring the characteristics $Z = \left(z_{\langle u,d\rangle}\right)_{u\in\mathcal{U},d\in\mathcal{D}}$. The complementarity function captures the non-z production increase from matches $\langle u_1, d_1\rangle$ and $\langle u_2, d_2\rangle$ from the baseline of the matches that arise when these two matched pairs exchange partners and the matches $\langle u_1, d_2\rangle$ and $\langle u_2, d_1\rangle$ arise.

Let a market-level type θ of complementarity functions be

$$\theta = \left(c\left(u_1, u_2, d_1, d_2, X\right), \ldots, c\left(u_{\mathcal{U}-1}, m_{\mathcal{U}}, w_{\mathcal{U}-1}, w_{\mathcal{U}}, X\right)\right) \tag{13}$$

as functions of X, where again $\mathcal{D} = \mathcal{U}$. Each market will have its own set of $\mathcal{U}^2\left(\mathcal{U}-1\right)^2/4$ heterogeneous complementarity functions. We will identify a distribution $G\left(\theta \mid X, Z\right)$ over complementarity functions formed by valid non-z production functions at the match level. We use the location normalization introduced in Section 3.3, where we set the non-z production of matches involving an upstream firm and a downstream firm with the same index to be 0, or

$$f_{\langle u,u\rangle}\left(x_u^{\text{up}}, x_u^{\text{down}}, x_{\langle u,u\rangle}\right) = 0, \forall u = 1, \ldots, \mathcal{U}.$$

The results that we will prove indicate that complementarity functions entirely characterize the outcome of the matching game and that the distribution $G(\theta \mid X, Z)$ of complementarity functions is identified. Therefore, the object being identified is sharp in the sense that it can be used to construct counterfactual matching probabilities, if say the elements of X change. As an intuitive check, notice that the examples of non-identification in Section 3.3 have the same values for complementarity functions for θ_1 and θ_2.

Lemma 3.5. *In a one-to-one matching game that allows only matched agents, the sum of non-z production functions in any assignment A can be expressed as a sum of non-z complementarity functions.*

Proof. Inside the proof, we suppress the arguments $\left(x_u, x_d, x_{\langle u,d\rangle}\right)$ of the non-z part of a production function (2). We prove the lemma by mathematical induction on $\mathcal{U}$, the number of upstream firms, which is equal to $\mathcal{D}$, the number of downstream firms. When $\mathcal{U} = 2$, the total non-z production in assignment $\{\langle 1,1\rangle, \langle 2,2\rangle\}$ is 0 by our location normalization, and the total non-z production in the other assignment $\{\langle 1,2\rangle, \langle 2,1\rangle\}$ is

$$f_{\langle 1,2\rangle} + f_{\langle 2,1\rangle} = -c\left(1, 2, 1, 2, X\right).$$

Thus the lemma is true for $\mathcal{U} = 2$. The induction hypothesis is that the lemma is true for $\mathcal{U} - 1$ for some arbitrary $\mathcal{U}$. For the case of markets with $\mathcal{U}$ upstream firms, there are two different types of assignments A that can occur. Recall that, in a given assignment A, we denote the downstream firm matched to upstream firm u as the value of the mapping $A(u)$.

Case 1. A is such that $A(\mathcal{U}) = \mathcal{U}$, where again $\mathcal{U}$ is the last upstream firm (and hence downstream firm) index. Then

$$\sum_{u=1}^{\mathcal{U}} f_{\langle u,A(u)\rangle} = \sum_{u=1}^{\mathcal{U}-1} f_{\langle u,A(u)\rangle} + f_{\langle \mathcal{U},\mathcal{U}\rangle} = \sum_{u=1}^{\mathcal{U}-1} f_{\langle u,A(u)\rangle} + 0,$$

by the location normalization. We can define an assignment A^* on the $(\mathcal{U}-1)$-by-$(\mathcal{U}-1)$ submarket involving the first $\mathcal{U}-1$ upstream firms and the first $\mathcal{U}-1$ downstream firms, such that $A^*(u) = A(u), \forall u < \mathcal{U}$, or all upstream firms other than the last upstream firm. Then

$$\sum_{u=1}^{\mathcal{U}} f_{\langle u,A(u)\rangle} = \sum_{u=1}^{\mathcal{U}-1} f_{\langle u,A(u)\rangle} + 0 = \sum_{u=1}^{\mathcal{U}-1} f_{\langle u,A^*(u)\rangle}.$$

By the induction hypothesis, we know that the last sum can be written as a sum of complementarity functions. Therefore the total non-z production given assignment A can be written as the sum of complementarity functions.

Case 2. A is such that $A(\mathcal{U}) = d^\star$, where $d^\star < \mathcal{U}$ as $\mathcal{U} = \mathcal{D}$ is the last index. Because all agents are matched, among upstream firms $1, \ldots, \mathcal{U}-1$ there exists an upstream firm $u^\star$ who matches with the last downstream firm $\mathcal{U}$: $A(u^\star) = \mathcal{U}$. Therefore

$$\begin{aligned}\sum_{u=1}^{\mathcal{U}} f_{\langle u,A(u)\rangle} &= \sum_{\substack{u=1,\\ u\neq u^\star}}^{\mathcal{U}-1} f_{\langle u,A(u)\rangle} + f_{\langle \mathcal{U},A(\mathcal{U})\rangle} + f_{\langle u^\star,A(u^\star))\rangle} \\ &= \sum_{\substack{u=1,\\ u\neq u^\star}}^{\mathcal{U}-1} f_{\langle u,A(u)\rangle} + f_{\langle \mathcal{U},d^\star\rangle} + f_{\langle u^\star,\mathcal{U})\rangle}. \quad (14)\end{aligned}$$

Recall $c\,(\mathcal{U}, u^\star, d^\star, \mathcal{U}, X) = f_{\langle \mathcal{U},d^\star\rangle} + f_{\langle u^\star,\mathcal{U}\rangle} - \left(f_{\langle \mathcal{U},\mathcal{U}\rangle} + f_{\langle u^\star,d^\star\rangle}\right)$. Because $f_{\langle \mathcal{U},\mathcal{U}\rangle} = 0$ by the location normalization, we have

$$f_{\langle \mathcal{U},d^\star\rangle} + f_{\langle u^\star,\mathcal{U}\rangle} = c\,(\mathcal{U}, u^\star, d^\star, \mathcal{U}, X) + f_{\langle u^\star,d^\star\rangle}.$$

Substitute this equation into (14):

$$\sum_{u=1}^{\mathcal{U}} f_{\langle u,A(u)\rangle} = f_{\langle 1,A(1)\rangle} + \ldots + f_{\langle u^\star,d^\star\rangle} + \ldots + f_{\langle \mathcal{U}-1,A(\mathcal{U}-1)\rangle} + c\left(\mathcal{U},u^\star,d^\star,\mathcal{U},X\right).$$

Notice that the first $(\mathcal{U}-1)$ terms only involve the first $(\mathcal{U}-1)$ upstream firms and the first $(\mathcal{U}-1)$ downstream firms. We can apply a similar argument to the one we used before to re-write the sum of these terms as the sum of non-z production from an assignment on a $(\mathcal{U}-1)$-by-$(\mathcal{U}-1)$ submarket. By the induction hypothesis, this is a sum of complementarity functions. Therefore $\sum_{u=1}^{\mathcal{U}} f_{\langle u,A(u)\rangle}$ is a sum of complementarity functions.

Thus the induction is complete, and the lemma is true for all $\mathcal{U}$. □

We have shown that the sums of production for assignments can be decomposed into the sums of complementarity functions. We next argue that two different vectors θ of complementarity functions, or the types of market-level unobservables, give a unique sum of non-z production for at least one assignment. For any vector θ of complementarity functions, we let f be a possibly non-unique vector of production functions that give the complementarity functions in θ.

Lemma 3.6. *Let θ_1 and θ_2 be two vectors of complementarity functions. θ_1 and θ_2 are equal if and only if the sums of non-z production functions of every assignment are the same across θ_1 and θ_2.*

Proof. Inside the proof, we suppress the arguments $\left(x_u, x_d, x_{\langle u,d\rangle}\right)$ of the non-z part of a production function (2). We prove the "if" part first. Denote the non-z complementarity functions and production functions for θ_1 as c^1, f^1 and for type θ_2 as c^2, f^2. Given any upstream firms u_1, u_2 and downstream firms d_1, d_2, we want to show that $c^1\left(u_1,u_2,d_1,d_2,X\right) = c^2\left(u_1,u_2,d_1,d_2,X\right)$ as functions of X. Consider assignment A_1 such that $A_1(u_1) = d_1, A_1(u_2) = d_2$, and the assignment A_2 such that $A_2(u_1) = d_2, A_2(u_2) = d_1$, where $A_1(u) = A_2(u)$ for all other upstream firms u. By the assumption that the sum of non-z production functions for each assignment is the same for both θ_1 and θ_2, we have by the hypothesis in the lemma that the sums of non-z production functions for the two types are equal for all assignments,

$$\sum_{u\neq u_1,u_2} f^1_{\langle u,A_1(u)\rangle} + f^1_{\langle u_1,A_1(u_1)\rangle} + f^1_{\langle u_2,A_1(u_2)\rangle} = \sum_{u\neq u_1,u_2} f^2_{\langle u,A_1(u)\rangle} + f^2_{\langle u_1,A_1(u_1)\rangle} + f^2_{\langle u_2,A_1(u_2)\rangle}$$
$$\sum_{u\neq u_1,u_2} f^1_{\langle u,A_2(u)\rangle} + f^1_{\langle u_1,A_2(u_1)\rangle} + f^1_{\langle u_2,A_2(u_2)\rangle} = \sum_{u\neq u_1,u_2} f^2_{\langle u,A_2(u)\rangle} + f^2_{\langle u_1,A_2(u_1)\rangle} + f^2_{\langle u_2,A_2(u_2)\rangle}.$$

Substitute in the values for A_1 and A_2, and subtract the equations:

$$f^1_{\langle u_1,d_1\rangle} + f^1_{\langle u_2,d_2\rangle} - \left(f^1_{\langle u_1,d_2\rangle} + f^1_{\langle u_2,d_1\rangle}\right) = f^2_{\langle u_1,d_1\rangle} + f^2_{\langle u_2,d_2\rangle} - \left(f^2_{\langle u_1,d_2\rangle} + f^2_{\langle u_2,d_1\rangle}\right),$$

which is precisely $c^1(u_1, u_2, d_1, d_2, X) = c^2(u_1, u_2, d_1, d_2, X)$.

The "only if" part is a direct result of Lemma 3.5. If the complementarity functions are the same, the sums of production functions of any assignment, which can be written as a sum of complementarity functions, must be the same across θ_1 and θ_2 as well. $\square$

We can use the previous lemma to show identification of the distribution of complementarity functions, which again by Lemma 3.5 is enough to compute assignments under many counterfactuals.

Theorem 3.7. *The distribution $G(\theta \mid X, Z)$ of the complementarity functions is identified in the one-to-one matching model where all agents are matched and the unobservable production functions are match-specific.*

Proof. Let the the vector of complementarity functions be (13) and let the vector of sums of non-z production of each assignment be

$$s(\theta, X) = \left(\sum_{u=1}^{\mathcal{U}} f_{\langle u, A_k(u)\rangle}\left(x_u, x_{A_k(u)}, x_{\langle u, A_k(u)\rangle}\right)\right)_{k=1}^{\mathcal{U}!},$$

where a one-to-one matching market with $\mathcal{U}$ upstream firms and $\mathcal{D} = \mathcal{U}$ downstream firms has $\mathcal{U}!$ assignments without any unmatched agents. Recall that the pairwise stable assignment for θ and (X, Z) maximizes the sum of production.

We show separability. We are given a finite set T of market-level types θ. If $\theta_1 \neq \theta_2$, then by the contrapositive of the "if" part of Lemma 3.6 and the properties of analytic functions, there exists an X such that $s(\theta_1, X) \neq s(\theta_2, X)$ for all types $\theta_1 \neq \theta_2$. Therefore, for T and X there exists a type θ_1 that corresponds to a minimal vector $s(\theta_1, X)$ with non-z production levels denoted as f. Let $z_{\langle u,d\rangle} = -f_{\langle u,d\rangle}\left(x_u, x_d, x_{\langle u,d\rangle}\right)$. For type θ_1, the total sum of production inclusive of z for each assignment is 0. By the minimal vector property, all other types in T have an assignment with total production strictly greater than 0, and this assignment is not the "diagonal" assignment $(\langle 1,1\rangle, \cdots, \langle \mathcal{U}, \mathcal{U}\rangle)$, whose non$-z$ production is always 0 by the location normalization. Now let $z_{\langle u,d\rangle}, \forall u \neq d$ decrease by a sufficiently small amount so that the pairwise stable assignment of all other types $\theta \neq \theta_1$ does not change. Then the payoff of the diagonal assignment is 0, and all other assignments has a slight negative payoff. Therefore θ is the only market-level type in T that gives the pairwise stable assignment $\{\langle 1,1\rangle, \cdots, \langle \mathcal{U}, \mathcal{U}\rangle\}$. The model is separable and hence $G(\theta \mid X, Z)$ is identified. $\square$

3.5 Matched Agents and Agent-Specific Unobservables

An important case of empirical interest is when the functions capturing primitive heterogeneity are agent-specific, reflecting agent preferences and unobserved characteristics. For markets with only matched agents, this is a special case of the previous section's analysis. However, the previous section assumed that each $z_{\langle u,d\rangle}$ is match-specific and lives in a product space. Many matching datasets will not have regressors that vary so flexibly at the match level. This section considers the case where z_u and z_d vary at the level of the agent and not the level of the match and we use data on only matched agents.

Let the non-z production function be of the same form as the product of index functions g_u and g_d in (11). As our model does not allow unmatched agents, we need additional normalizations. We first normalize $g_1^{\text{up}}(x_1) = 0$ for $u = 1$ and for all X, by subtracting an equal amount from all $g_u^{\text{up}}(x_u)$. This location normalization does not affect the outcome: the sum of non-z production for a given assignment A is decreased by a constant times the sum of the $g_d(x_d)$, which is independent of A. Therefore, the ordering of the total, non-z production for the different assignments does not change. Similarly, we normalize $g_1^{\text{down}} = 0$. Next, we apply a scale normalization and set $g_2^{\text{up}} = 1$. This is equivalent to setting $\tilde{g}_u^{\text{up}} = \frac{g_u^{\text{up}}}{g_2^{\text{up}}}$, and $\tilde{g}_d^{\text{down}} = g_d^{\text{down}} \cdot g_2^{\text{up}}$. This change does not alter the assignment level of production for every assignment, because for each match, the production is still $\tilde{g}_u^{\text{up}} \cdot \tilde{g}_d^{\text{down}} = g_u^{\text{up}} \cdot g_d^{\text{down}}$. Then for $\mathcal{D} = \mathcal{U}$, we define a market-level type to be (adopting the newly normalized values of g's and dropping the tilde):

$$\theta = \left((g_u)_{u=3}^{\mathcal{U}}, (g_d)_{d=2}^{\mathcal{D}} \right). \tag{15}$$

In other words, θ is comprised of $\mathcal{U} + \mathcal{D} - 3$, again for $\mathcal{D} = \mathcal{U}$, functions of x_u or x_d. We show that the distribution of θ is identified.

Let θ be a market-level type. We let $w(\theta, X)$ be the matrix of non-z production levels

$$w(\theta, X) = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & f_{\langle 2,2\rangle}(x_2, x_2) & \dots & f_{\langle 2,\mathcal{U}\rangle}(x_2, x_{\mathcal{U}}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f_{\langle \mathcal{U},2\rangle}(x_{\mathcal{U}}, x_2) & \dots & f_{\langle \mathcal{U},\mathcal{U}\rangle}(x_{\mathcal{U}}, x_{\mathcal{U}}) \end{pmatrix},$$

where we define

$$f_{\langle u,d\rangle}(x_u, x_d) = g_u(x_u) \cdot g_d(x_d).$$

We first show that if the sums of production for all assignments are the same across θ_1 and θ_2, then the non-z production of all matches $g_u(x_u) \cdot g_d(x_d)$ must be the same for θ_1 and θ_2.

Lemma 3.8. *Let θ_1 and θ_2 be two market-level types* (15)*. If the sums of non-z production for all assignments are the same for θ_1 and θ_2, then $\theta_1 = \theta_2$ as functions of X.*

Proof. Again, in terms of the notation we suppress the arguments to a production function. Let f^1 index the non-z production functions for type θ_1 and let f^2 index the non-z production functions for type θ_2. We start on the diagonal elements: for assignment $(\langle 1,1\rangle, \langle 2,2\rangle, \cdots, \langle \mathcal{U},\mathcal{U}\rangle)$, we have

$$\sum_{u=2}^{\mathcal{U}} f^1_{\langle u,u\rangle} = \sum_{u=2}^{\mathcal{U}} f^2_{\langle u,u\rangle}, \tag{16}$$

as the missing first term in each sum is 0 as $g_1^{\text{up}}(x_1)$ is normalized to be 0 for all X. For assignment $(\langle 1,k\rangle, \langle 2,2\rangle \cdots, \langle k,1\rangle, \cdots, \langle \mathcal{U},\mathcal{U}\rangle)$, we have

$$\sum_{\substack{u=2\\ u\neq k}}^{\mathcal{U}} f^1_{\langle u,u\rangle} = \sum_{\substack{u=2\\ u\neq k}}^{\mathcal{U}} f^2_{\langle u.u\rangle}, \tag{17}$$

again as the matches $\langle 1,k\rangle$ and $\langle k,1\rangle$ have zero non-z production by the location normalization that $g_1(x_1) = 0$ for upstream firms and $g_1(x_1) = 0$ for downstream firms. Subtract equation (17) from equation (16), which gives us $f^1_{\langle k,k\rangle} = f^2_{\langle k,k\rangle}$. The previous argument can be repeated for all k, giving $f^1_{\langle k,k\rangle} = f^2_{\langle k,k\rangle}\, \forall\, k = 1, \ldots, \mathcal{U}$.

Next we induct on the size of the $K \times K$ upper left submatrix of $w(\theta, X)$. For the upper left two-by-two matrix, the only nonzero term is $f_{\langle 2,2\rangle}$, which we just showed was identical between θ_1 and θ_2. Now assume the $(\mathcal{U}-1)$-by-$(\mathcal{U}-1)$ upper left submatrix of $w(\theta, X)$ is the same between θ_1 and θ_2. We know the element $f_{\langle \mathcal{U},\mathcal{U}\rangle}$ is the same. For arbitrary elements $f_{\langle \mathcal{U},d\rangle}$of the last row and $f_{\langle u,\mathcal{U}\rangle}$ of the last column, we want to show that those elements are identical between θ_1 and θ_2. If we can show this for the an element of the last row, a symmetric argument will hold for an element of the last column. Focusing on $f_{\langle \mathcal{U},d\rangle}$, consider the assignment

$$(\langle \mathcal{U}-1,1\rangle, \langle \mathcal{U}-2,2\rangle, \cdots, \langle \mathcal{U}-d+1,d-1\rangle, \langle \mathcal{U},d\rangle, \langle \mathcal{U}-d,d+1\rangle, \cdots, \langle d,\mathcal{U}\rangle),$$

which has the sum of non-z production

$$0 + f_{\langle \mathcal{U}-2,2\rangle} + \cdots + f_{\langle \mathcal{U}-d+1,d-1\rangle} + f_{\langle \mathcal{U},d\rangle} + f_{\langle \mathcal{U}-d,d+1\rangle} + \cdots + 0.$$

All terms in the sum except for $f_{\langle \mathcal{U},d\rangle}$ are equal across θ_1 and θ_2 by the induction hypothesis and the location normalization. Because the sums of non-z production are equal across assignments, $f^1_{\langle \mathcal{U},d\rangle} = f^2_{\langle \mathcal{U},d\rangle}, \forall d$. Thus the induction is complete and $w(\theta_1, X) = w(\theta_2, X)$.

With the normalization that $g_2^{\text{up}} = 1$, and $f_{\langle u,d\rangle}(x_u, x_d) = g_u(x_u) \cdot g_d(x_d)$, we have $g_2(x_u) \cdot g^1_d(x_d) = g_2(x_u) \cdot g^2_d(x_d)$, which implies that $g^1_d(x_d) = g^2_d(x_d)$. By comparing every term that involves $f_{\langle u,d\rangle}, u \neq 1, 2$, we have $g^1_u(x_u) = g^2_u(x_u)$. Therefore $\theta_1 = \theta_2$ □

The social planner can be seen as a single agent making discrete choices between assignments. It is easy to see that the difference between the production levels of assignments governs the pairwise stable assignment that will occur. We now show that different θ's give varying differences in the production of assignments.

Lemma 3.9. *Let A_1 be an assignment. Then given types θ_1 and θ_2 such that $\theta_1 \neq \theta_2$, there exists an assignment A_2 such that*

$$\sum_{u=1}^{\mathcal{U}} f^1_{\langle u,A_1(u)\rangle}\left(x_u, x_{A_1(u)}\right) - \sum_{u=1}^{\mathcal{U}} f^1_{\langle u,A_2(u)\rangle}\left(x_u, x_{A_2(u)}\right) \neq \\ \sum_{u=1}^{\mathcal{U}} f^2_{\langle u,A_1(u)\rangle}\left(x_u, x_{A_1(u)}\right) - \sum_{,u=1}^{\mathcal{U}} f^2_{\langle u,A_2(u)\rangle}\left(x_u, x_{A_2(u)}\right) \quad (18)$$

as functions of X.

Proof. We drop the arguments of non-z production functions inside the proof. We prove the lemma by contradiction. Assume (18) is always an equality for all A_2 and X. Then by re-arranging terms, we have

$$\sum_{u=1}^{\mathcal{U}} f^1_{\langle u,A_2(u)\rangle} - \sum_{u=1}^{\mathcal{U}} f^2_{\langle u,A_2(y)\rangle} = \sum_{u=1}^{\mathcal{U}} f^1_{\langle u,A_1(u)\rangle} - \sum_{u=1}^{\mathcal{U}} f^2_{\langle u,A_1(u)\rangle}$$

as functions of X. We can define $D(X)$ to be equal to the value of the right side, which is not a function of A_2. Consider the assignment $A_2 = (\langle 1,1\rangle, \langle 2,2\rangle \cdots, \langle k,k\rangle, \cdots, \langle \mathcal{U},\mathcal{U}\rangle)$. Given the implication of the location normalizations that $f_{\langle 1,1\rangle} = 0$, we have

$$\sum_{u=2}^{\mathcal{U}} f^1_{\langle u,u\rangle} - \sum_{u=2}^{\mathcal{U}} f^2_{\langle u,u\rangle} = D(X).$$

Consider the assignment $A_3 = (\langle 1,k\rangle, \langle 2,2\rangle \cdots, \langle k,1\rangle, \cdots, \langle \mathcal{U},\mathcal{U}\rangle)$. Under the implication of the location normalizations that $f_{\langle 1,k\rangle} = 0$, we have

$$\sum_{\substack{u=2 \\ u\neq k}}^{\mathcal{U}} f^1_{\langle u,u\rangle} - \sum_{\substack{u=2 \\ u\neq k}}^{\mathcal{U}} f^2_{\langle u,u\rangle} = D(X).$$

Comparing the equations for A_2 and A_3 and then varying the argument for each k, we have that $f^1_{\langle k,k\rangle} = f^2_{\langle k,k\rangle} \forall k$ as functions of X. Therefore $\sum_{u=2}^{\mathcal{U}} f^1_{\langle u,u\rangle} = \sum_{u=2}^{\mathcal{U}} f^2_{\langle u,u\rangle}$ and $D(X) = 0$. Therefore the sums of non-z production for each assignment are the same across the two types for each X. By Lemma 3.8, this implies that $\theta_1 = \theta_2$, and we have a contradiction. $\square$

Theorem 3.10. *The distribution $G(\theta \mid X, Z)$ is identified in the one-to-one matching model with agent-specific characteristics, agent-specific unobservable index functions, and without unmatched agents.*

Proof. We suppress the arguments of production functions and index functions. We prove separability. We are given a finite set T of different market-level types θ. Fix an assignment A_0 as a reference assignment. Consider the vector of differences of sums production between the other $\mathcal{U}!-1$ assignments (there are no unmatched agents) and A_0:

$$d(\theta, X) = \left(\sum_{u=1}^{\mathcal{U}} f_{u,A_1(u)} - \sum_{u=1}^{\mathcal{U}} f_{u,A_0(u)}, \cdots, \sum_{u=1}^{\mathcal{U}} f_{u,A_{\mathcal{U}!-1}(u)} - \sum_{u=1}^{\mathcal{U}} f_{u,A_0(u)} \right).$$

By a property of analytic functions and Lemma 3.9, there exists X such that $d(\theta_1, X) \neq d(\theta_2, X)$, if $\theta_1 \neq \theta_2$. Then there exists a minimal vector of type θ_1. Recall $f_{\langle u,d \rangle} = g_u \cdot g_d$. Let $z_u = -g_u$ and $z_d = g_d$ for type θ_1. Then the total payoffs of each assignment is 0 for type θ_1. For all other types, there exists at least one assignment different from A_0 and whose payoff is strictly greater than that of A_0. Therefore the equilibrium outcomes of all types in T but θ_1 are different from A_0. Type θ_1 has the same production between all types, and we can trivially move z_u and z_d to break this indifference without affecting the outcomes of other types. Thus, there is a singleton type θ_1 that has pairwise stable assignment A_0 at (X, Z) and the model is separable and hence identified. □

4 Many-to-Many Matching

4.1 Notation for Many-to-Many Matching

Many-to-many matching models of business to business relationships are common in industrial organization and related fields such as marketing and strategy. Think about the example of upstream and downstream firms, studied in Fox (2010a). In that case, one upstream firm can supply many downstream firms and one downstream firm can be supplied by many upstream firms. A **match group** centered around the upstream firm u is $\langle u, D_u \rangle$, where $D_u = \{d_1, d_2, \cdots, d_k\}$ is the **partner list** of k downstream firms matched with upstream firms u. A set of full matches centered around the upstream firm u is $\left\{ \langle u, d, t_{\langle u,d \rangle} \rangle \right\}_{d \in D_u}$, where, as before, $t_{\langle u,d \rangle}$ is the transfer between u and d. The transfer is not in the data.

Agents such as u and d in a many-to-many matching game have integer quotas q_u and q_d, which can be infinity in some applications. Upstream firm u can only have as many physical matches as q_u, although in a particular assignment u may have less than q_u matches. We will assume that quotas are in the data. Fox (2010b) uses a different identification strategy that allows quotas to not be known. Let Q collect the quotas for all firms in a market. Q can be a non-trivial random variable or it can be the same across markets. To avoid repeating the definition of identification, we simply add Q to the vector of non-z observables X. Concepts such as outcomes, assignments and pairwise stability generalize straightforwardly to the case of many-to-many matching. The existence and uniqueness of a pairwise stability assignment is dependent on further assumptions, as not all specifications of production in a many-to-many matching game make pairwise stability equal to full stability and hence give existence and uniqueness with probability 1.

For upstream firms, we generalize the definition of profit in (1) to

$$v^{\text{up}}_{\langle u,D_u\rangle}\left(x^{\text{up}}_u,\left(x^{\text{down}}_d\right)_{d\in D_u},\left(x_{\langle u,d\rangle}\right)_{d\in D_u},x_{\langle u,D_u\rangle}\right)+\gamma^{\text{up}}z_{\langle u,D_u\rangle}+\sum_{d\in D_u}t_{\langle u,d\rangle},$$

where the vector $x_{\langle u,D_u\rangle}$ and scalar $z_{\langle u,D_u\rangle}$ are observable characteristics that vary flexibly at the level of the match group $\langle u, D_u\rangle$. We also consider the case of using firm characteristics z_u and z_d in place of $z_{\langle u,d\rangle}$ below. We allow the non-z profit of an upstream to not be additively separable across multiple matches, or that

$$v^{\text{up}}_{\langle u,D_u\rangle}\left(x^{\text{up}}_u,\left(x^{\text{down}}_d\right)_{d\in D_u},\left(x_{\langle u,d\rangle}\right)_{d\in D_u},x_{\langle u,D_u\rangle}\right)\neq\sum_{d\in D_u}v^{\text{up}}_{\langle u,\{d\}\rangle}\left(x^{\text{up}}_u,x^{\text{down}}_d,x_{\langle u,d\rangle}\right),$$

although we can allow for additive separability across multiple matches as a special case. In order to define a production function, we must assume that downstream firm profits are additively separable across multiple upstream firms, or that the total profit of downstream firm d satisfies

$$v^{\text{down}}_{\langle U_d,d\rangle}\left(\left(x^{\text{up}}_u\right)_{u\in U_d},x^{\text{down}}_d,\left(x_{\langle u,d\rangle}\right)_{u\in U_d}\right)+\sum_{d\in D_u}t_{\langle u,d\rangle}=\sum_{u\in U_d}v^{\text{down}}_{\langle\{u\},d\rangle}\left(x^{\text{up}}_u,x^{\text{down}}_d,x_{\langle u,d\rangle}\right)-\sum_{d\in D_u}t_{\langle u,d\rangle}.$$

Because of this additivity, the term $z_{\langle u,D_u\rangle}$ specific to the match group $\langle u, D_u\rangle$ can only enter upstream firm u's profits. With the restriction to additive separability across multiple matches for downstream firms only, total production from the match group $\langle u, D_u\rangle$ becomes

$$v^{\text{up}}_{\langle u,D_u\rangle}\left(x^{\text{up}}_u,\left(x^{\text{down}}_d\right)_{d\in D_u},\left(x_{\langle u,d\rangle}\right)_{d\in D_u},x_{\langle u,D_u\rangle}\right)+\sum_{u\in U_d}v^{\text{down}}_{\langle\{u\},d\rangle}\left(x^{\text{up}}_u,x^{\text{down}}_d,x_{\langle u,d\rangle}\right)+\gamma^{\text{up}}z_{\langle u,D_u\rangle}.$$

There are no covariates of the form $x_{\langle u,D_u\rangle}$. As before, we normalize $\gamma^{\text{up}}=1$ in order to impose monotonicity, as in the literature on multinomial choice (Ichimura and Thompson, 1998; Berry and Haile, 2010b; Fox and Gandhi, 2010). We also define the production function for a given match group

$\langle u, D_u \rangle$ to satisfy

$$f_{\langle u,D_u\rangle}\left(x_u^{\text{up}}, \left(x_d^{\text{down}}\right)_{d\in D_u}, \left(x_{\langle u,d\rangle}\right)_{d\in D_u}, x_{\langle u,D_u\rangle}\right) + z_{\langle u,D_u\rangle} =$$
$$v_{\langle u,D_u\rangle}^{\text{up}}\left(x_u^{\text{up}}, \left(x_d^{\text{down}}\right)_{d\in D_u}, \left(x_{\langle u,d\rangle}\right)_{d\in D_u}, x_{\langle u,D_u\rangle}\right) + \sum_{u\in U_d} v_{\langle\{u\},d\rangle}^{\text{down}}\left(x_u^{\text{up}}, x_d^{\text{down}}, x_{\langle u,d\rangle}\right) + z_{\langle u,D_u\rangle}. \quad (19)$$

We will seek to identify the distribution $G(\theta \mid X, Z)$ of market-level unobservables

$$\theta = \left(f_{\langle u,D_u\rangle}\right)_{u\in\mathcal{U}, D_u\subseteq\mathcal{D}, |D_u|\leq q_u}, \quad (20)$$

where quotas again are in X. As we have done with $\mathcal{U}$ and $\mathcal{D}$, we will overload notation and let D_u also be the number $|D_u|$ of downstream firms in the set D_u.

We could not define a production function without restricting the production of downstream firms to be additively separable across multiple upstream firm partners. In the future, we will explore relaxing this and separately identifying aspects of $v_{\langle u,D_u\rangle}^{\text{up}}$ and $v_{\langle U_d,d\rangle}^{\text{down}}$, which is done parametrically in Fox (2010a).

4.2 Identification Under a Full Stability Condition

Full stability is a strong equilibrium concept. Hatfield, Kominers, Nichifor, Ostrovsky and Westkamp (2011) study a relatively general class of matching games that nests many-to-many matching. They in part prove that any pairwise stable equilibrium will be fully stable whenever the profits of upstream firms and downstream firms satisfy conditions known as **full substitutes**. See their paper as well as Fujishige and Yang (2003) for discussion of this restrictive condition. Under full substitutes, the properties of existence of a pairwise stable outcome, uniqueness with probability 1 of the pairwise stable assignment and the ability to compute the assignment using a social planner's problem arise.

For the purposes of identification, a fully stable assignment is robust to deviations by coalitions of many firms, so the model imposes many restrictions on the data that help us identify distributions of functions $f_{\langle u,D_u\rangle}$ that take on the arguments of many firms. We will briefly discuss models that are pairwise stable but not fully stable below.

The results using full stability for many-to-many matching are extensions of the discussion for one-to-one matching.

4.3 Unmatched Firms, Match Group Characteristics in Z

Let production be of the form (19) and so the market-level unobservable θ is of the form (20)

Theorem 4.1. ***(Analog of Theorem 3.1)*** *Consider a many-to-many matching game with unmatched firms and where $z_{\langle u,D_u\rangle}$ varies at the match group level. The distribution $G(\theta \mid X, Z)$ is* *identified if the outcome is fully stable.*

We omit the proof as it follows naturally from the proof of Theorem 3.1, which just sets the total production of all match groups to 0 for one type corresponding to a minimal vector.

4.4 Unmatched Firms, Firm Characteristics in X, Z

The previous result identifies a lot, but at the expenses of flexible variation in the data: $z_{\langle u,D_u\rangle}$ varies at the match group level. Many datasets will not have this flexible variation, although Fox (2010a) does. This section considers characteristics z_u and z_d as well as x_u and x_d only. Let the production to a match involving upstream firm u and a set of downstream firms $D_u = \{d_1, \ldots, d_k\}$ be

$$g_u(x_u) \prod_{d \in D_u} g_d(x_d) + z_u \prod_{d \in D_u} z_d, \tag{21}$$

where $g_u(x_u)$ is a heterogeneous index function for the upstream firm u, and g^d is such an index function for downstream firm d. The consequence of using less flexible variation in z is that the functional form is much more restrictive. Therefore, a market-level type is

$$\theta = \left((g_u)_{u \in \mathcal{U}}, (g_d)_{d \in \mathcal{D}} \right).$$

The proof of Theorem 3.2 only needs slight modification: $w(\theta, X)$, the vector of all production functions consists of multiplications of all index functions of firms in each match, and we set $z_u = -g_u$ and $z_d = g_d$ for all $u \in \mathcal{U}$ and $d \in \mathcal{D}$ after locating the minimal vector.

Theorem 4.2. ***(Analog of Theorem 3.2)*** *Consider a many-to-many matching game with unmatched firms and where observed characteristics vary at the firm level. The distribution $G(\theta \mid X, Z)$ is identified if the outcome is fully stable.*

4.5 All Quotas Filled, Match Group Characteristics in Z

We return to having data on characteristics specific to the match group, $z_{\langle u,D_u\rangle}$. However, we work with a matching model where all firms are required to fulfill their quotas. This somewhat extreme model is the natural extension of the cases in one-to-one matching without unmatched agents. Otherwise, the decision to leave a quota slot unfilled introduces an aspect of individual rationality, which lets the proof techniques in Theorems 3.1 and 4.1 be used. By requiring that all firms fill their quotas, we use only the sorting pattern for identification and so work with the hardest possible case. For it to be physically feasible for all quotas to be filled, we require the sum of quotas of downstream firms equal the sum of quotas of upstream firms, or $\sum_{d \in \mathcal{D}} q_d = \sum_{u \in \mathcal{U}} q_u$.

Following the result in Theorem 3.7, we need to extend the definition of a **complementarity function**. Consider matches $\langle u, D_u\rangle$ and $\langle \tilde{u}, D_{\tilde{u}}\rangle$. Also consider $d \in D_u, \tilde{d} \in D_{\tilde{u}}$, such that $d \notin D_{\tilde{u}}$

and $\tilde{d} \notin D_u$. Let $D'_u = (D_u \backslash \{d\}) \cup \left\{\tilde{d}\right\}$ and $D'_{\tilde{u}} = \left(D_{\tilde{u}} \backslash \left\{\tilde{d}\right\}\right) \cup \{d\}$. The complementarity function is

$$c\left(u, D_u, D'_u, \tilde{u}, D_{\tilde{u}}, D'_{\tilde{u}}\right) = f_{\langle u, D_u\rangle}\left(x_u, (x_d)_{d\in D_u}, \left(x_{\langle u,d\rangle}\right)_{d\in D_u}, x_{\langle u, D_u\rangle}\right) + f_{\langle \tilde{u}, D_{\tilde{u}}\rangle}\left(x_{\tilde{u}}, (x_d)_{d\in D_{\tilde{u}}}, \left(x_{\langle \tilde{u},d\rangle}\right)_{d\in D_{\tilde{u}}}, x_{\langle \tilde{u}, D_{\tilde{u}}\rangle}\right) - f_{\langle u, D'_u\rangle}\left(x_u, (x_d)_{d\in D'_u}, \left(x_{\langle u,d\rangle}\right)_{d\in D'_u}, x_{\langle u, D'_u\rangle}\right) - f_{\langle \tilde{u}, D'_{\tilde{u}}\rangle}\left(x_{\tilde{u}}, (x_d)_{d\in D'_{\tilde{u}}}, \left(x_{\langle \tilde{u},d\rangle}\right)_{d\in D'_{\tilde{u}}}, x_{\left\langle \tilde{u}, D'_{\tilde{u}}\right\rangle}\right).$$

As in one-to-one matching, we notationally write the complementarity function as having firm indices for arguments instead of observed characteristics, for conciseness. The complementarity function is the difference in the sums of two production functions, when *one* downstream firm from a partner list D_u switches with a different downstream firm in another partner list $D_{\tilde{u}}$. We normalize $f\left(u, A_0\left(u\right)\right) = 0\,\forall u$ for some assignment A_0, where $A\left(u\right)$ returns the partner list D_u of upstream firm u under assignment A. Therefore, a market-level unobservable θ is

$$\theta = \left(c\left(u, D_u, D'_u, \tilde{u}, D_{\tilde{u}}, D'_{\tilde{u}}\right)\right) \text{ for } u \in \mathcal{U}, D_u \subseteq \mathcal{D}, d \in D_u, \tilde{u} \in \mathcal{U}, D_{\tilde{u}} \subseteq \mathcal{D}, \tilde{d} \in D_{\tilde{u}}, \quad D'_u = (D_u \backslash \{d\}) \cup \left\{\tilde{d}\right\}, D'_{\tilde{u}} = \left(D_{\tilde{u}} \backslash \left\{\tilde{d}\right\}\right) \cup \{d\}$$

In Lemma 3.5 for one-to-one matching, we showed that the sum of production for each assignment is equal to a sum of complementarity functions. This result also has a many-to-many counterpart. This means the complementarity functions are sufficient for computing counterfactuals.

Lemma 4.3. ***(Analog of Lemma 3.5)*** *Total production for an assignment in many-to-many matching games can be written as a sum of complementarity functions, provided that in all assignments the quotas of all agents are known and filled.*

Proof. We drop the arguments of production functions inside the proof. The intuition is that any given assignment $\tilde{A}$ is transformed to the normalized to A_0 after a finite steps of exchanges of two downstream firms between two match groups. Then the sum of production for $\tilde{A}$ can be constructed as the series of complementarity functions associated with the swaps. Focusing on two downstream firms at a time, suppose $\tilde{d} \in \tilde{A}\left(u_1\right)$, $d \in \tilde{A}\left(u_2\right)$, $d \notin A_0\left(u_2\right)$, and $\tilde{d} \notin A_0\left(u_1\right)$. Starting from $\tilde{A}$ and moving to a particular new assignment A_2,

$$\mathcal{C}\left(u_1,\tilde{A}\left(u_1\right),A_2\left(u_1\right),u_2,\tilde{A}\left(u_2\right),A_2\left(u_2\right)\right)=$$
$$f_{\left\langle u_1,\tilde{A}(u_1)\right\rangle}+f_{\left\langle u_2,\tilde{A}(u_2)\right\rangle}-\left(f_{\left\langle u_1,\left(\tilde{A}(u_1)\backslash\tilde{d}\right)\cup\{d\}\right\rangle}+f_{\left\langle u_2,\left(\tilde{A}(u_2)\backslash d\right)\cup\left\{\tilde{d}\right\}\right\rangle}\right)=$$
$$\sum_{u=1}^{\mathcal{U}}f_{\left\langle u,\tilde{A}(u)\right\rangle}-\sum_{u=1}^{\mathcal{U}}f_{\left\langle u,A_2(u)\right\rangle},$$

where A_2 is defined to the assignment such that $A_2\left(u_1\right)=\left(\tilde{A}\left(u_1\right)\backslash\tilde{d}\right)\cup\{d\}$, $A_2\left(u_2\right)=\left(\tilde{A}\left(u_2\right)\backslash d\right)\cup\left\{\tilde{d}\right\}$, and $A_2\left(u\right)=\tilde{A}\left(u\right)$ for other upstream firms. Following this pattern for other pairs of downstream firms, we can construct A_3 from A_2 and so forth until all downstream firms matched to different upstream firms in $\tilde{A}$ and A_0 are exchanged and $A_K=A_0$, the assignment whose production is normalized to 0, for some K. The sum of production for $A_K=A_0$ differs from that of $\tilde{A}$ by a sum of complementarity functions. Because the sum of production for A_0 is 0, the sum of production for $\tilde{A}$ is a sum of production functions. □

Note that the proof implicitly uses induction on the number of upstream firms, similarly to Lemma 3.5. The two proofs are similar. We can also extend our result for one-to-one matching that two market-level types θ, based on complementarity functions, lead to assignments whose sums of production vary across θ_1 and θ_2. Recall this is interesting because the differences in assignment production govern which assignment is pairwise stable.

Lemma 4.4. ***(Analog of Lemma 3.6)*** *Let θ_1 and θ_2 be two vectors of complementarity functions. θ_1 and θ_2 are equal if and only if the sums of non-z production functions of every assignment are the same across θ_1 and θ_2.*

Proof. Inside the proof, we drop the arguments of production functions. First, prove the "if" part by contradiction, so $\theta_1\neq\theta_2$. As $\theta_1\neq\theta_2$, at least one complementarity function is different, or there exists u, $\tilde{u}$, … such that

$$f^1_{\langle u,D_u\rangle}+f^1_{\langle\tilde{u},D_{\tilde{u}}\rangle}-f^1_{\langle u,D'_u\rangle}-f^1_{\langle\tilde{u},D'_{\tilde{u}}\rangle}\neq f^2_{\langle u,D_u\rangle}+f^2_{\langle\tilde{u},D_{\tilde{u}}\rangle}-f^2_{\langle u,D'_u\rangle}-f^2_{\langle\tilde{u},D'_{\tilde{u}}\rangle}.$$

Following the argument in Lemma 3.6, simply form assignments A_1 and A_2 that are identical except for the exchanges of two downstream firms between u and $\tilde{u}$. This is a contradiction to the sums of non-z production being the same. The "only if" part is a direct result of Lemma 4.3. □

With this lemma, we can show identification.

Theorem 4.5. ***(Analog of Theorem 3.7)*** *Consider a many-to-many matching game where all quota slots are filled and where observed characteristics vary at the match group level. The distribution*

$G(\theta \mid X, Z)$ of the complementarity functions is identified if the outcome is always fully stable.

Given the lemmas, the proof is analogous to Theorem 3.7 and is therefore omitted.

4.6 All Quotas Filled, Firm Characteristics in X, Z

Lemmas 4.3 and 4.4 do not involve the elements of Z at all and so also cover the agent-specific case when the production to a match group is given by (21). Therefore, we can identify the distribution of complementarity functions in the firm-specific Z case as long as we can set the production of all matches for a given type θ to be zero, which we can under the production function in (21).

Theorem 4.6. ***(Analog to the case of Theorem 3.10, results differ)*** *Consider a many-to-many matching game where all quota slots are filled and where observed characteristics vary at the match group level. The distribution $G(\theta \mid X, Z)$ of the complementarity functions is identified if the outcome is always fully stable.*

The identification results under full stability nest Theorem 3.1, Theorem 3.2 and Theorem 3.7. However, we are only able to identify the distribution of complementarity functions in the agent-specific case in many-to-many matching, whereas the distribution of the agent-specific unobservable is identified in one-to-one matching.

4.7 Using Pairwise Stability Only

In car parts application in Fox (2010a), the full substitutes condition in Hatfield et al. (2011) is not satisfied, so pairwise stable outcomes will not be automatically fully stable. Further, there are hundreds of upstream and downstream firms, and so imposing full stability as a primitive solution concept requires an unrealistically large volume of communication among firms to form coalitions that maximize total production. To approximate reality, Fox followed much of the matching literature and worked with the implications of a weaker solution concept, pairwise stability. A deviation under the notion of pairwise stability involves at most two firms communicating (hence the name “pairwise”). Therefore the volume of communication is significantly reduced.

There are two new issues that arise when relying on only pairwise stability, instead of full stability. First, without the full substitutes condition, a pairwise stable (and hence a fully stable) outcome might not exist. In other words, for some combination of X, Z and θ, there may exist no pairwise stable assignments. This is analogous to the non-existence of Nash equilibria when attention is restricted to pure strategies, for noncooperative games. Second, without full stability, a pairwise stable assignments will often not be unique. This is analogous to multiplicity of pure strategy Nash equilibria. These issues are subjects of our ongoing research.

5 Conclusions

Matching models that have been structurally estimated to date have not allowed rich models of unobservables. It has been an open question whether data on who matches with whom as well as agent characteristics are enough to identify distributions of unobservables. In this paper, we explore several sets of conditions that lead to identification.

One lesson is that having regressors that vary at the match and not the agent level helps identify unobservables that nest both unobserved agent characteristics and unobserved agent preferences for observed characteristics. Therefore, matching models that have restricted attention to homogeneous preferences as functions of observable characteristics and additive, match-specific errors that are i.i.d. across matches are not making full use of the data. One can allow for this richness when working with matching data. Alternatively, characteristics that vary only at the firm level can help identify distributions of unobserved characteristics, but not unobserved preferences, as unobserved characteristics are nested in our index functions, while general unobserved preferences for partner characteristics are not.

There are important differences in the results using data on both matched and unmatched firms and the results requiring that all firms be matched. In the former case, relatively simple proofs show that general primitive distributions of heterogeneity are identified. Using data on only matched firms, one can identify distributions of what we call complementarity functions. We introduce this concept and show that complementarity functions determine the sum of production for assignments, so complementarity functions are enough to calculate assignment probabilities under counterfactuals.

A Proof of Theorem 2.3: Separability Implies Identification

The proof is by contradiction. Say the model is separable but the distribution $G(\theta \mid X, Z)$ is not identified. Then there exist $G^1(\theta \mid X, Z)$ and $G^2(\theta \mid X, Z)$ such that they both generate the same assignment probabilities $\Pr(A \mid X, Z)$. These distributions have finite supports T^1 and T^2, respectively. Let $T^3 = T^1 \cup T^2$ and extend G^1 and G^2 to T^3 by having say G^1give 0 weight to points in $T^3 - T^1$. Then based on (8)

$$\Pr\left(A \mid X, Z;\, G^1\right) - \Pr\left(A \mid X, Z;\, G^2\right) = \sum_{n=1}^{N} 1\left[A\,\text{pairwise stable assignment} \mid X, Z, \theta_n\right]\left(p_n^1(X, Z) - p_n^2(X, Z)\right) = 0\,\forall\, A, X, Z,$$

where $p_n^1(X, Z)$ is the weight given to θ_n under G^1.

Let T^4 be the types $\theta \in T^3$ such that $p_n^1(X,Z) \neq p_n^2(X,Z)$ as functions of X, Z. Then

$$\Pr\left(A \mid X, Z; G^1\right) - \Pr\left(A \mid X, Z; G^2\right) = \sum_{\theta_n \in T^4} 1\left[A \,\text{pairwise stable assignment} \mid X, Z, \theta_n\right]\left(p_n^1(X,Z) - p_n^2(X,Z)\right) = 0 \,\forall\, A, X, Z. \quad (22)$$

Apply separability to the finite set of market-level unobservables T^4. Then there exists $A^\star$, $X^\star$ and $Z^\star$ such that only type $\theta_{n^\star} \in T^4$ gives assignment $A^\star$ at $X^\star$ and $Z^\star$. So, at these $A^\star$, $X^\star$ and $Z^\star$,

$$\Pr\left(A^\star \mid X^\star, Z^\star; G^1\right) - \Pr\left(A^\star \mid X^\star, Z^\star; G^2\right) = p_n^1(X^\star, Z^\star) - p_n^2(X^\star, Z^\star),$$

where $p_n^1(X,Z)$ and $p_n^2(X,Z)$ are not equal as functions of X and Z, by the definition of T^4. Because of Assumption 2.2.2, there exists an $(X^{\star\star}, Z^{\star\star})$ where only type $\theta_{n^\star} \in T^4$ gives assignment $A^\star$ and $p_n^1(X^{\star\star}, Z^{\star\star}) \neq p_n^2(X^\star, Z^\star)$. This is a contradiction to (22) and so the distribution $G(\theta \mid X, Z)$ is identified.

References

Ackerberg, Dan, Jinyong Hahn, and Geert Ridder, "Testable Implications of Multiple Unobservables in Structural Nonparametric Nonseparable Models," 2010. working paper.

Ackerberg, Daniel A. and Maristella Botticini, "Endogenous Matching and the Empirical Determinants of Contract Form," *Journal of Political Economy*, 2002, *110* (2002), 564–591.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, July-August 1973, *81* (4), 813–846.

Berry, Steven and Philip A. Haile, "Identification in Differentiated Products Markets Using Market Level Data," February 2010. working paper.

Berry, Steven T., "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 1992, *60* (4), 889–917.

_ **and Philip A. Haile**, "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," 2010. Yale University working paper.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Analyzing the Determinants of the Matching Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets," 2003. University of Buffalo working paper.

Boyd, JH and RE Mellman, "Effect of Fuel Economy Standards on the U. S. Automotive Market: An Hedonic Demand Analysis," *TRANSPORT. RES.*, 1980, *14* (5), 367–378.

Bresnahan, Timothy F. and Peter C. Reiss, "Entry and Competition in Concentrated Markets," *The Journal of Political Economy*, October 1991, *99* (5), 977–1009.

Cardell, NS and FC Dunbar, "Measuring the societal impacts of automobile downsizing," *TRANSPORT. RES.*, 1980, *14* (5), 423–434.

Chen, Jiawei, "Two-Sided Matching and Spread Determinants in the Loan Market," May 2009. working paper.

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

Dagsvik, John K., "Aggregation in Matching Markets," *International Economic Review*, February 2000, *41* (1), 27–57.

Echenique, F., S.M. Lee, and M. Shum, "Aggregate matchings," 2010. Working paper.

Fox, Jeremy T., "Estimating Matching Games with Transfers," 2010. University of Michigan working paper.

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

_ **and Amit Gandhi**, "Nonparametric Identification and Estimation of Random Coefficients in Nonlinear Economic Models," 2010. University of Michigan working paper.

Fujishige, S. and Z. Yang, "A note on Kelso and Crawford's gross substitutes condition," *Mathematics of Operations Research*, 2003, *28* (3), 463–469.

Gordon, Nora and Brian Knight, "A spatial merger estimator with an application to school district consolidation," *Journal of Public Economics*, 2009, *93*, 752–765.

Hatfield, John William, Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp, "Stability and Competitive Equilibrium in Trading Networks," January 2011. Stanford University working paper.

Hitsch, Günter, Ali Hortaçsu, and Dan Ariely, "Matching and Sorting in Online Dating," *American Economic Review*, 2010.

Ichimura, H. and TS Thompson, "Maximum likelihood estimation of a binary choice model with random coefficients of unknown distribution," *Journal of Econometrics*, 1998, *86* (2), 269–295.

Koopmans, Tjalling C. and Martin Beckmann, "Assignment Problems and the Location of Economic Activities," *Econometrica*, January 1957, *25* (1), 53–76.

Levine, Anna A., "Licensing and Scale Economies in the Biotechnology Pharmaceutical Industry," April 2009. Harvard University working paper.

Manski, Charles F., "Maximum Score Estimation of the Stochastic Utility Model of Choice," *Journal of Econometrics*, 1975, *3* (3), 205–228.

Ostrovsky, Michael, "Stability in Supply Chain Networks," *American Economic Review*, December 2008, *98* (3), 897–923.

Park, Minjung, "An Empirical Two-Sided Matching Model of Acquisitions: Understanding Merger Incentives and Outcomes in the Mutual Fund Industry," September 2008. working paper.

Roth, Alvin E. and Marilda A. Oliveira Sotomayor, *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* Econometric Society Monographs, Cambridge University Press, 1990.

Shapley, Lloyd S. and Martin Shubik, "The assignment game I: the core," *International Journal of Game Theory*, 1972, *1*, 111–130.

Sørensen, Morten, "How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital," *Journal of Finance*, December 2007, *LXII* (6), 2725–2762.

Yang, Yupin, Mengze Shi, and Avi Goldfarb, "Estimating the Value of Brand Alliances in Professional Team Sports," *Marketing Science*, November-December 2009, *28* (6), 1095–1111.