

NBER WORKING PAPER SERIES

UNOBSERVED HETEROGENEITY IN MATCHING GAMES

Jeremy T. Fox  
Chenyu Yang

Working Paper 18168  
<http://www.nber.org/papers/w18168>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2012

Thanks to Stephan Lauermann, Ariel Pakes, and seminar participants at various conferences and universities for helpful suggestions. Our email addresses are [jtfox@umich.edu](mailto:jtfox@umich.edu) and [chnyyang@umich.edu](mailto:chnyyang@umich.edu). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Jeremy T. Fox and Chenyu Yang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Unobserved Heterogeneity in Matching Games

Jeremy T. Fox and Chenyu Yang

NBER Working Paper No. 18168

June 2012

JEL No. C35,C78,J12,L0

**ABSTRACT**

Agents in two-sided matching games vary in characteristics that are unobservable in typical data on matching markets. We investigate the identification of the distribution of these unobserved characteristics using data on who matches with whom. The distribution of match-specific unobservables cannot be fully recovered without information on unmatched agents, but the distribution of a combination of unobservables, which we call unobserved complementarities, can be identified. Knowledge of the unobserved complementarities is sufficient to construct certain counterfactuals. The distribution of agent-specific unobservables is identified under different conditions.

Jeremy T. Fox  
Economics Department  
University of Michigan  
238 Lorch Hall  
611 Tappan Ave  
Ann Arbor, MI 48109  
and NBER  
jeremyfox@gmail.com

Chenyu Yang  
Economics Department  
University of Michigan  
238 Lorch Hall  
611 Tappan Ave  
Ann Arbor, MI 48109  
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. This paper is partially motivated by such applications in industrial organization, where downstream firms pay upstream firms money, and thus it is reasonable to work with transferable utility matching games (Koopmans and Beckmann, 1957; Becker, 1973; Shapley and Shubik, 1972).

There has been recent interest in the structural estimation of (both transferable utility and non-transferable utility) matching games.<sup>1</sup> The papers we cite are unified in estimating some aspect of the preferences of agents in a matching game from data on who matches with whom as well as the observed characteristics of agents or of matches. The sorting patterns 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 discrete, non-cooperative (Nash) games, like the entry literature in industrial organization (Bresnahan and Reiss, 1991; Berry, 1992; Mazzeo, 2002; Tamer, 2003; Seim, 2006; Bajari et al., 2010) and the discrete outcomes peer effects literature (Brock and Durlauf, 2007; de Paula and Tang, 2012). Matching games typically use the cooperative solution concept of pairwise stability.<sup>2</sup>

The empirical literature cited previously structurally estimates how various structural or equilibrium objects, such as payoffs or preferences, are functions of the characteristics of agents *observed in the data*. For example, Choo and Siow (2006) study the marriage market in the United States and estimate how the equilibrium payoffs of men for women vary by the ages of the man and the woman. Sørensen (2007) studies the matching of venture capitalists to entrepreneurs as a function of observed venture capitalist experience. 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 agent profit functions.

The above papers all use data on a relatively limited set of agent characteristics. In Choo and Siow, personality and physical attractiveness are not measured, even though those characteristics are likely important in determining the equilibrium pattern of marriages. Similarly, in Fox each firm’s product quality is not directly measured and is only indirectly inferred. In Sørensen, the unobserved ability

---

<sup>1</sup>See, among others: Dagsvik (2000); Boyd et al. (2003); Choo and Siow (2006); Sørensen (2007); Fox (2010a); Gordon and Knight (2009); Chen (2009); Ho (2009); Park (2008); Yang et al. (2009); Logan et al. (2008); Levine (2009); Baccara et al. (forthcoming); Siow (2009); Galichon and Salanie (2010); Chiappori et al. (2010b); Crawford and Yurokoglou (forthcoming); Weese (2010); Christakis et al. (2010); Echenique et al. (2011); Menzel (2011).

<sup>2</sup>Transferable utility matching games can be seen as special cases of models of hedonic equilibrium (Brown and Rosen, 1982; Ekeland et al., 2004; Heckman et al., 2010; Chiappori et al., 2010a). Unlike the empirical literature on hedonic equilibrium, the estimation approaches in most matching papers do not rely on data on equilibrium prices or transfers. Compared to the current work, the hedonic papers do not allow for unobserved characteristics.

of each venture capitalist is not measured. 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. Akerberg and Botticini (2002) provide empirical evidence that farmers and landlords sort on unobservables such as risk aversion and monitoring ability, without formally estimating a matching game or the distribution of these unobservables.

Our discussion of the empirical applications cited above shows that unobserved characteristics are potentially important. As the consistency of estimation procedures for matching games depends on assumptions on the unobservables, empirical conclusions might be more robust if the estimated matching games allow richly specified distributions of unobserved agent heterogeneity. This paper investigates what data on the sorting patterns between agents can tell us about the distributions of unobserved agent characteristics relevant for sorting. In particular, we study the nonparametric identification of distributions of unobserved agent heterogeneity in two-sided matching games. With the distribution of unobservables, the researcher can explain sorting and construct counterfactual predictions about market assignments. This paper allows for this empirically relevant heterogeneity in partner preferences using data on only observed matches (who matches with whom), not data from, say, an online dating site on rejected profiles (Hitsch et al., 2010) or on equilibrium transfers, such as wages in a labor market (Eeckhout and Kircher, 2011). Transfers are often confidential data in firm contracts (Fox, 2010a) and are rarely observed in marriage data (Becker, 1973).

In the following specific sense, this paper on identification is ahead of the empirical matching literature because no empirical papers have parametrically estimated distributions of unobserved characteristics in matching games. Thus, this paper seeks to introduce a new topic for economic investigation, rather than to simply loosen parametric restrictions in an existing empirical literature. This paper contributes to the literature on the nonparametric (allowing infinite dimensional objects) identification of transferable utility matching games (Fox, 2010b; Graham, 2010). Our paper is distinguished because of its focus on identifying distributions of unobservables, rather than mostly deterministic functions of observables. Our focus on using data on many markets with finite numbers of agents in each (transferable utility) market follows Fox (2010a,b).<sup>3</sup>

We first consider a baseline model, which is stripped down to focus on the key problem of identifying distributions of heterogeneity from sorting data. In our baseline transferable utility matching game,

---

<sup>3</sup>In addition to our study of identifying distributions of unobservables, there are many modeling differences between our paper and the literature on transferable utility matching games following the approach of Choo and Siow (2006), including Galichon and Salanie (2010) and Chiappori et al. (2010b). We use data on many markets with finite numbers of players and different realizations of observables and unobservables in each market; the Choo and Siow approach applies to one large market with an infinity of agents. We require at least one continuous, observable characteristic per match or per agent; the Choo and Siow literature allows only a finite number of observable characteristic values. Unobservables in the Choo and Siow literature are typically i.i.d. shocks for the finite observable types rather than than unobserved agent characteristics or unobserved preferences on observed, ordered characteristics, such as random coefficients. This so-called “separability” assumption in the Choo and Siow literature has the key implication that the unobservable, economically endogenous transfers in the Choo and Siow literature are independently distributed (within the single market they consider) from the unobservable, economically exogenous econometric errors. The unobserved transfers in our model are not independently distributed from the unobserved agent characteristics, for example.

the primitive that governs sorting is the matrix that collects the production values for each potential match in a matching market. The production level of each match is additively separable in observable and unobservable terms. The observable term is captured by a match-specific regressor. The unknown primitive is therefore the distribution (representing randomness across markets) of the matrix that collects the unobservable terms in the production of each match in a market. We call this distribution the distribution of match-specific unobservables. Match-specific unobservables nest many special cases, such as agent-specific unobservables.

We provide three main results, and some extensions. Our first main result states that the distribution of match-specific unobservables is not identified in a one-to-one matching game with data on who matches with whom but without data on unmatched or single agents. Our second main result states that the distribution of a change of variables of the unobservables, the distribution of what we call unobserved complementarities, is identified. We precisely define unobserved complementarities below. Our identification proof works by tracing the joint (across possible matches in a market) cumulative distribution function of these unobserved complementarities using the match-specific observables. We also show that knowledge of the distribution of unobserved complementarities is sufficient for computing assignment probabilities. Our third main result says that the distribution of the primitively specified, match-specific unobservables is actually identified when unmatched agents are observed in the data.

Our three main results can be intuitively understood by reference to a classic result in Becker (1973). He studies sorting in two-sided, transferable utility matching games where agents have scalar characteristics (types). He shows that high-type agents match to high-type agents if the types of agents are complements in the production of matches. Many production functions for match output exhibit complementarities. Say in Becker’s model male and female types are  $x_m$  and  $x_w$ , respectively. A production function with horizontal preferences, such as  $-(x_m - x_w)^2$ , and one with vertical preferences, such as  $2x_mx_w$ , can both have the same cross-partial derivative, here 2. Becker’s result that complementarities alone drive sorting means that data on sorting cannot tell these two production functions apart. In our more general class of matching games, our first main result is that we cannot identify the distribution of match-specific unobservables. Our second main result is that we can identify the distribution of our notion of unobserved complementarities. These two results are analogous to Becker’s results for a more general class of matching games.

Our third main result uses data on unmatched agents. In a matching game, agents can unilaterally decide to be single or not. If all other agents are single and hence available to match, the fact that one particular agent is single can only be explained by the production of all matches involving that agent being less than the production from being single. This type of direct comparison between the production of being single and the production of being matched is analogous to the way identification proceeds in discrete Nash games, where the payoff of a player’s observed (in the data) strategy must be higher than strategies not chosen, given the strategies of rivals. Thus, the availability of data on

unmatched agents introduces an element of individual rationality that maps directly into the data and is therefore useful for identification of the primitive distribution of match-specific characteristics.

Many empirical researchers might be tempted to specify a parametric distribution of match-specific unobservables. Our three results together suggest that estimating a matching model with a parametric distribution of match-specific unobservables will not lead to credible estimates without using data on unmatched agents, as a more general nonparametrically specified distribution is not identified. One could instead impose a parametric distribution for unobserved complementarities.

We examine several extensions to the baseline model that add more empirical realism. Our baseline model imposes additive separability between unobservables and observables in the production of a match. We examine an extension where additional observed characteristics enter match production and these characteristics may, for example, have random coefficients on them, reflecting the *random preferences* of agents for partner characteristics. For example, observationally identical men are often observed to marry observationally distinct women. One important hypothesis is that these men have heterogeneous preferences for the observable characteristics of women. In a model with random preferences, we identify the distribution of match production conditional on the characteristics of agents and matches other than the match-specific characteristics used in the baseline model. This object of identification follows identification work using special regressors in the multinomial choice literature (Lewbel, 2000; Matzkin, 2007; Berry and Haile, 2010).

In another extension, we identify fixed-across-markets but heterogeneous-within-a-market coefficients on the the match-specific characteristics used in the baseline model. This relaxes the assumption that the match-specific characteristics enter the production of each match in the same manner. Another extension considers models where key observables vary at the agent and not the match level. We can achieve identification of the primitive distribution of match-specific unobservables without relying on data on unmatched agents, but by imposing a perhaps stronger functional form for match production.

Our results on one-to-one, two-sided matching games extend naturally to one-sided matching. An example of a one-sided matching problem is mergers between firms. Our results also extend to many-to-many matching under a strong restriction on preferences known as substitutes, which rules out multiple pairwise stable assignments occurring with positive probability. We briefly discuss the literature on identification under multiple equilibria in Nash games, but combining approaches to multiple equilibria with matching games is outside the scope of our paper.

## 2 Baseline Identification Results

We mainly analyze a two-sided, one-to-one matching game with transferable utility (Koopmans and Beckmann, 1957; Becker, 1973; Shapley and Shubik, 1972; Roth and Sotomayor, 1990, Chapter 8). This section imposes that all agents must be matched in order to focus purely on the identification coming from

agent sorting and not from the individual rationality decision to be single. We also use a simple covariate space. We change these assumptions in later sections.

## 2.1 Baseline Model

We will use the terms “agents” and “firms” interchangeably. In a one-to-one matching game, an upstream firm  $u$  matches with a downstream firm  $d$ . Upstream firm  $u$  and downstream firm  $d$  can form a **physical match**  $\langle u, d \rangle$ . The monetary transfer from  $d$  to  $u$  is denoted as  $t_{\langle u, d \rangle}$ ; we will not require data on the transfers. In a solution to the game,  $u$  and  $d$  may form a **full match**  $\langle u, d, t_{\langle u, d \rangle} \rangle$ . The **production** from a match  $\langle u, d \rangle$  is

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

where  $z_{\langle u, d \rangle}$  is a regressor specific to match  $\langle u, d \rangle$  and  $e_{\langle u, d \rangle}$  is a match-specific unobservable. An example of a match-specific regressor  $z_{\langle u, d \rangle}$  is the distance between the headquarters of firms  $u$  and  $d$ ; we discuss agent-specific regressors in later sections. We can more primitively model production for a match  $\langle u, d \rangle$  as the sum of the profit of  $u$  and the profit of  $d$ , where the transfer  $t_{\langle u, d \rangle}$  between  $d$  and  $u$  enters additively separably into both profits and therefore cancels in their sum.<sup>4</sup> Only production levels matter for the matches that form, and we will not attempt to identify upstream firm profits separately from downstream firm profits.

There are  $N$  firms on each side of the market, and in this section there can be no single firms.  $N$  can also represent the set  $\{1, \dots, N\}$ . The matrix

$$\begin{pmatrix} z_{\langle 1, 1 \rangle} + e_{\langle 1, 1 \rangle} & \cdots & z_{\langle 1, N \rangle} + e_{\langle 1, N \rangle} \\ \vdots & \ddots & \vdots \\ z_{\langle N, 1 \rangle} + e_{\langle N, 1 \rangle} & \cdots & z_{\langle N, N \rangle} + e_{\langle N, N \rangle} \end{pmatrix}$$

describes the production of all matches in a market, where the rows are upstream firms and the columns are downstream firms. Let

$$E = \begin{pmatrix} e_{\langle 1, 1 \rangle} & \cdots & e_{\langle 1, N \rangle} \\ \vdots & \ddots & \vdots \\ e_{\langle N, 1 \rangle} & \cdots & e_{\langle N, N \rangle} \end{pmatrix}, Z = \begin{pmatrix} z_{\langle 1, 1 \rangle} & \cdots & z_{\langle 1, N \rangle} \\ \vdots & \ddots & \vdots \\ z_{\langle N, 1 \rangle} & \cdots & z_{\langle N, N \rangle} \end{pmatrix}$$

be the matrices of unobservables and observables, respectively, in a market.<sup>5</sup>

<sup>4</sup>If the profit of  $u$  at some market outcome is  $\pi_{\langle u, d \rangle}^u + t_{\langle u, d \rangle}$  and the profit of  $d$  is  $\pi_{\langle u, d \rangle}^d - t_{\langle u, d \rangle}$ , then the production of the match  $\langle u, d \rangle$  is equal to  $\pi_{\langle u, d \rangle}^u + \pi_{\langle u, d \rangle}^d = z_{\langle u, d \rangle} + e_{\langle u, d \rangle}$ . We will not attempt to learn the distributions of the unobservable portions of  $\pi_{\langle u, d \rangle}^u$  and  $\pi_{\langle u, d \rangle}^d$  separately (Fox, 2010b).

<sup>5</sup>Because the scalar  $z_{\langle u, d \rangle}$  is an element of the matrix  $Z$ , we do not use upper and lower case letters (or other notation)

A feasible one-to-one **assignment**  $A$  is a set of physical matches  $A = \{\langle u_1, d_1 \rangle, \dots, \langle u_N, d_N \rangle\}$ , where for this section each firm is matched exactly once. An **outcome** is a list of full matches:

$$\{\langle u_1, d_1, t_{\langle u_1, d_1 \rangle} \rangle, \dots, \langle u_N, d_N, t_{\langle u_N, d_N \rangle} \rangle\}.$$

An outcome is **pairwise stable** if it is robust to deviations by pairs of two firms, as defined in references such as Roth and Sotomayor (1990, Chapter 8).<sup>6</sup> An assignment  $A$  is called **pairwise stable** if there exists an underlying outcome that is pairwise stable. The literature cited previously proves that the existence of a pairwise stable assignment is guaranteed and that an assignment  $A$  is pairwise stable if and only if it maximizes the sum of production

$$S(A, E, Z) = \sum_{\langle u, d \rangle \in A} (z_{\langle u, d \rangle} + e_{\langle u, d \rangle}).$$

If  $z_{\langle u, d \rangle}$  or  $e_{\langle u, d \rangle}$  have continuous support,  $S(A, E, Z)$  has a unique maximizer with probability 1 and therefore the pairwise stable assignment is unique with probability 1. The sum of the **unobserved production** is likewise

$$\tilde{S}(A, E) = \sum_{\langle u, d \rangle \in A} e_{\langle u, d \rangle}.$$

A market is defined to be the pair  $(E, Z)$ ; agents in a market can match and agents in different markets cannot. A researcher observes the assignment  $A$  and the regressors  $Z$  for many markets. In other words, in each matching market we observe who matches with whom  $A$  and the characteristics  $Z$  of the realized and potential matches. This allows the identification of  $\Pr(A | Z)$ . Researchers do not observe  $t_{\langle u, d \rangle}$ , which is usually part of confidential contracts.

$Z$  is independent of the unobservable matrix  $E$ .  $E$  has the joint cumulative distribution function (CDF)  $G(E)$ . We assume that  $Z$  has full and product support, meaning that any  $Z \in \mathbb{R}^{N^2}$  is observed. Each element of  $Z$  is called a **special regressor**, as each  $z_{\langle u, d \rangle}$  enters production with an additive functional form, the sign and coefficient on each  $z_{\langle u, d \rangle}$  in production is common across matches (normalized to be 1), each  $z_{\langle u, d \rangle}$  has large support, and  $Z$  is independent of  $E$ . We allow a match-specific coefficient on each  $z_{\langle u, d \rangle}$  and, separately, use only agent-specific regressors in later extensions of the baseline model. Consider the matching of patients to medical providers. The distance from each patient's house to each medical provider is an observable, match-specific characteristic.<sup>7</sup>

Such special regressors have been used to prove point identification in the binary and multinomial choice literature (Manski, 1988; Ichimura and Thompson, 1998; Lewbel, 1998, 2000; Matzkin, 2007;

---

to distinguish random variables and their realizations. Whether we refer to a random variable or its realization should be clear from context.

<sup>6</sup>We omit standard definitions here that can be easily found in the literature.

<sup>7</sup>Distance  $z_{\langle u, d \rangle}$  is always positive and likely enters match production with a negative sign; we can always construct a new regressor  $\tilde{z}_{\langle u, d \rangle} = -(z_{\langle u, d \rangle} - E[z_{\langle u, d \rangle}])$  that enters with a positive sign and has mean zero.



Gautier and Kitamura, 2011; Berry and Haile, 2010; Fox and Gandhi, 2012). In this literature, failure to have a special regressor often results in set rather than point identification of the distribution of heterogeneity.<sup>8</sup> We discuss the failure of the support condition in our matching context in Section 2.5. In this paper, we use special regressors in part to focus on reasons specific to matching games for the failure of point identification.<sup>9</sup>

## 2.2 Data Generating Process and Identification

The unknown primitive whose identification we first explore is the CDF  $G(E)$ , which reflects how the unobservables vary across matching markets and hence (because we do not assume independence across the  $e_{\langle u,d \rangle}$ 's) within matching markets as well. By not restricting  $G$ , we allow for many special cases, such as  $e_{\langle u,d \rangle} = e_u \cdot e_d$ , where  $e_u$  is a scalar unobserved upstream firm characteristic and  $e_d$  is a scalar unobserved downstream firm characteristic. We do not restrict the support of  $G(E)$  except in normalizations that follow.

The probability of assignment  $A$  occurring given the observables  $Z$  is

$$\Pr(A | Z; G) = \int_E 1[A \text{ pairwise stable assignment} | Z, E] dG(E), \quad (1)$$

where  $1[A \text{ pairwise stable assignment} | Z, E]$  is equal to 1 when  $A$  is a pairwise stable assignment for the market  $(E, Z)$ .

The distribution  $G$  is said to be **identified** whenever, for  $G^1 \neq G^2$ ,  $\Pr(A | Z; G^1) \neq \Pr(A | Z; G^2)$  for some pair  $(A, Z)$ .  $G^1$  and  $G^2$  give a different probability for at least one assignment  $A$  given  $Z$ . If  $G$  has continuous and full support so that all probabilities  $\Pr(A | Z; G)$  are nonzero (for every  $A$ ,  $S(A, E, Z)$  will be maximized by a range of  $E$ ) and continuous in the elements of  $Z$ , the existence of one such pair  $(A, Z)$  implies that a set of  $Z$  with positive measure satisfies  $\Pr(A | Z; G^1) \neq \Pr(A | Z; G^2)$ .

All but one of our identification results will be constructive, in that we can trace a distribution such as  $G(E)$  using variation in an object such as  $Z$ . Regardless, all identification arguments can be used to prove the consistency of a nonparametric mixtures estimator for a distribution such as  $G$ , as

---

<sup>8</sup>Consider a binary choice model of buying a can of soda (or not) where the special regressor is the (negative) price of the soda, which varies across the dataset. If we assume that consumers' willingnesses to pay for the can of soda are bounded by \$0 and \$10, we can point identify the distribution of the willingness to pay for soda if observable prices range between \$0 and \$10. If prices range only between \$0 and \$5, we can identify the fraction of consumers with values above \$5 by seeing the fraction who purchase at \$5. We cannot identify the fraction with values above \$6, or any value greater than \$5. If we assume nothing about the support of the willingness to pay, we need prices to vary across all of  $\mathbb{R}$  (including negative prices if consumers may have negative willingnesses to pay) for point identification of the distribution of the willingness to pay for soda.

<sup>9</sup>Our use of large support and the use of large support in most of the literature on binary and multinomial choice does not constitute identification at infinity as used to study Nash games by, for example, Tamer (2003). Identification at infinity in the Nash games setting uses only extreme values of regressors for all but one player to, in effect, turn a multi-player game into a single-player decision problem. We use large regressor values only to identify the tails of distributions of heterogeneity.

Fox and Kim (2011) show for a particular, computationally simple mixtures estimator.<sup>10</sup> Other mixtures estimators can be used, including simulated maximum likelihood, the EM algorithm, NPMLE, and MCMC.<sup>11</sup>

As maximizing  $S(A, E, Z)$  determines the assignment seen in the data, the ordering of  $S(A, E, Z)$  across assignments  $A$  as a function of  $E$  and  $Z$  is a key input to identification. We can add a constant to the production of all matches involving the same upstream firm and the ordering of the production  $S(A, E, Z)$  of all assignments will remain the same. Therefore, we impose the location normalization that sets the production of all matches  $\langle i, i \rangle$  to 0, or

$$E = \begin{pmatrix} 0 & e_{\langle 1,2 \rangle} & \cdots & e_{\langle 1,N \rangle} \\ e_{\langle 2,1 \rangle} & 0 & \cdots & e_{\langle 2,N \rangle} \\ \vdots & \vdots & \ddots & \vdots \\ e_{\langle N,1 \rangle} & e_{\langle N,2 \rangle} & \cdots & 0 \end{pmatrix}. \quad (2)$$

The need for a location normalization of this sort is already a non-identification result: we cannot identify whether the production levels of all matches involving one firm are higher than the production levels of all matches involving a second firm. This non-identification result is unsurprising: the differential production of matches and hence assignments governs the identity of the pairwise stable assignment in any market.

### 2.3 Non-Identification

We will show another non-identification result. Consider the two realizations of matrices of unobservables

$$E_1 = \begin{pmatrix} 0 & e_{\langle 1,2 \rangle} & \cdots & e_{\langle 1,N \rangle} \\ e_{\langle 2,1 \rangle} & 0 & \cdots & e_{\langle 2,N \rangle} \\ \vdots & \vdots & \ddots & \vdots \\ e_{\langle N,1 \rangle} & e_{\langle N,2 \rangle} & \cdots & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & e_{\langle 1,2 \rangle} + 1 & \cdots & e_{\langle 1,N \rangle} \\ e_{\langle 2,1 \rangle} - 1 & 0 & \cdots & e_{\langle 2,N \rangle} - 1 \\ \vdots & \vdots & \ddots & \vdots \\ e_{\langle N,1 \rangle} & e_{\langle N,2 \rangle} + 1 & \cdots & 0 \end{pmatrix}.$$

It is easy to verify that  $S(A, E_1, Z) = S(A, E_2, Z)$  for all  $A, Z$ , which means that the optimal assignment  $A$  is the same for  $E_1$  and  $E_2$ , for any  $Z$ . Therefore it is not possible to separately identify the relative frequencies of  $E_1$  and  $E_2$  in the data generating process. We summarize the

<sup>10</sup>The proof of consistency in Fox and Kim (2011) requires the random variable (such as  $E$ ) to have compact support, which is not required here for identification.

<sup>11</sup>For large markets, these estimators all have computational problems arising from the combinatorics underlying the set of matching game assignments. Fox (2010a) uses a maximum score estimator to avoid these computational problems, but does not estimate a distribution of unobservables. Our identification arguments do not address computational issues.

counterexample in the following non-identification theorem.

**Theorem 1.** *The distribution  $G(E)$  of market-level unobserved match characteristics is **not** identified in a matching game where all agents must be matched.*

Consider a simple case focusing on two upstream firms and two downstream firms. If we see the matches  $\langle u_1, d_1 \rangle$  and  $\langle u_2, d_2 \rangle$  in the data, we cannot know whether this assignment forms because  $\langle u_1, d_1 \rangle$  has high production,  $\langle u_2, d_2 \rangle$  has high production,  $\langle u_1, d_2 \rangle$  has low production, or  $\langle u_2, d_1 \rangle$  has low production. The non-identification result implies that parametric estimation of  $G(E)$  under these assumptions may not be well founded, in that the generalization removing the parametric restrictions is not identified.

## 2.4 Unobserved Complementarities

As described in the introduction, Becker (1973) shows that complementarities govern sorting when there is one characteristic (schooling) per agent. Likewise, while it is not possible to identify the distribution of the most primitive unobserved heterogeneity, we will show that the distribution of unobserved complementarities, defined below, can be identified.

**Definition.** The **unobserved complementarity** between upstream firms  $u_1, u_2$  and downstream firms  $d_1, d_2$  is defined to be

$$c(u_1, u_2, d_1, d_2) = e_{\langle u_1, d_1 \rangle} + e_{\langle u_2, d_2 \rangle} - (e_{\langle u_1, d_2 \rangle} + e_{\langle u_2, d_1 \rangle}). \quad (3)$$

The unobserved complementarities capture the change in the unobserved production when two matched pairs  $\langle u_1, d_1 \rangle$  and  $\langle u_2, d_2 \rangle$  exchange partners and the matches  $\langle u_1, d_2 \rangle$  and  $\langle u_2, d_1 \rangle$  arise. As an intuitive check that the counterexample in the previous section will not prevent identification, notice that  $E_1$  and  $E_2$  have the same values for all unobserved complementarities.

The market-level vector comprising all unobserved complementarities is

$$C = (c(u_1, u_2, d_1, d_2) \mid u_1, u_2, d_1, d_2 \in N). \quad (4)$$

Given  $E$ , one can construct  $C$ . The location normalization in (2) on the underlying match-specific unobservables  $E$  is still used. We will prove that the unobserved complementarities characterize the outcome of the matching game and that the CDF  $F(C)$  of unobserved complementarities  $C$  is identified.

There are  $N^4$  values  $c(u_1, u_2, d_1, d_2)$  given any realization  $E$ . However, all unobserved complementarities can be formed from a smaller set of other unobserved complementarities by addition and subtraction.

**Lemma 1.** *There is a random vector*

$$B = (c(u_1, u_2, d_1, d_2) \mid u_1 = d_1 = 1, u_2, d_2 \in \{2, \dots, N\})$$

*of  $(N - 1)^2$  unobserved complementarities such that any unobserved complementarity  $c(u_1, u_2, d_1, d_2)$  in  $C$  is equal to a  $(u_1, u_2, d_1, d_2)$ -specific sum and difference of terms in  $B$ . The indices  $(u'_1, u'_2, d'_1, d'_2)$  of the terms in  $B$  in the sum do not depend on the realization of  $E$ .*

The proof is given in the appendix. We do not prove that  $B$  is a minimal vector, and it certainly is not a unique vector with these properties. The lemma provides the insight that identifying  $F(C)$  will not require the same dimension of moments in the data (i.e.  $\Pr(A \mid Z)$  for choices of  $A$  and  $Z$ ) as the dimension of  $F(C)$ , where  $C$  is a random vector of  $N^4$  elements. We will write about identifying the distribution  $F(C)$ , but just as well our goal could be identifying the distribution  $F_B(B)$ , as any realization of  $B$  implies a realization of  $C$ , and vice-versa. For parametric and nonparametric estimation using a finite sample, we recommend estimating  $F_B(B)$ . We should be clear that Lemma 1 provides important context but is not referenced in the proofs of other results.

Recall that  $\tilde{S}(A, E) = \sum_{\langle u, d \rangle \in A} e_{\langle u, d \rangle}$  is the unobserved production from assignment  $A$ . We show that knowing  $C$  implies knowing the sum  $\tilde{S}(A, E)$  of unobserved production for all assignments.

**Lemma 2.** *For each  $A$ ,  $\tilde{S}(A, E)$  is equal to an  $A$ -specific sum and difference of unobserved complementarities in  $C$ . The indices  $(u_1, u_2, d_1, d_2)$  of the terms in the sum do not depend on the realization of  $E$ .*

The proof is in the appendix. This lemma indicates that we can compute counterfactual assignment probabilities  $\Pr(A \mid Z; F)$  if we know the distribution  $F$  of unobserved complementarities  $C$ . Also by this lemma, we use the overloaded notation  $\tilde{S}(A, C)$  for the sum of unobserved production as a function of  $C$  instead of  $E$ .

We next argue that two different vectors  $C_1$  and  $C_2$  of unobserved complementarities give distinct sums of unobserved production for at least one assignment.

**Lemma 3.** *Consider two realizations  $C_1$  and  $C_2$  of the random vector  $C$ .  $C_1 = C_2$  if and only if  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all assignments  $A$ .*

By the lemma, identifying the distribution of  $C$  only requires identifying the distribution of  $\tilde{S}$ . This lemma is the key economic result that shows that there is hope for the identification of  $F(C)$ , as different realizations of  $C$  lead to different sums of unobserved production of assignments, which possibly lead to different assignments occurring in the data.

For examples, we verify the conclusions of Lemmas 1, 2 and 3 for the cases of  $N = 2$  and  $N = 3$ .

**Example 1.** Consider the case of  $N = 2$ . A matrix of match-specific unobservables is

$$E = \begin{pmatrix} 0 & e_{\langle 1,2 \rangle} \\ e_{\langle 2,1 \rangle} & 0 \end{pmatrix}.$$

There are two possible assignments,  $A_1 = \{\langle 1,1 \rangle, \langle 2,2 \rangle\}$  and  $A_2 = \{\langle 1,2 \rangle, \langle 2,1 \rangle\}$ . There is one random variable in the random vector  $B$ :  $c(1,2,1,2) = 0 + 0 - e_{\langle 1,2 \rangle} - e_{\langle 2,1 \rangle}$ . As for the other unobserved complementarity in  $C$ ,  $c(2,1,2,1) = -c(1,2,1,2)$ , which demonstrates Lemma 1. The sum of unobserved production for  $A_1$  is  $\tilde{S}(A_1, E) = 0$  by the location normalization (2). Also,  $\tilde{S}(A_2, E) = e_{\langle 1,2 \rangle} + e_{\langle 2,1 \rangle} = -c(1,2,1,2)$ . These formulas for  $\tilde{S}(A, E)$  for both assignments demonstrate Lemma 2. Now consider two realizations of the random vector  $C$ , namely  $C_1$  and  $C_2$ . As  $\tilde{S}(A_2, C) = -c(1,2,1,2)$ , it follows that  $C_1 = C_2$  if and only if  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all assignments  $A$ . This demonstrates Lemma 3.

**Example 2.** Consider the case of  $N = 3$ . A matrix of match-specific unobservables is

$$E = \begin{pmatrix} 0 & e_{\langle 1,2 \rangle} & e_{\langle 1,3 \rangle} \\ e_{\langle 2,1 \rangle} & 0 & e_{\langle 2,3 \rangle} \\ e_{\langle 3,1 \rangle} & e_{\langle 3,2 \rangle} & 0 \end{pmatrix}.$$

There are six possible assignments,

$$\begin{aligned} A_1 &= \{\langle 1,1 \rangle, \langle 2,2 \rangle, \langle 3,3 \rangle\} \\ A_2 &= \{\langle 1,2 \rangle, \langle 2,1 \rangle, \langle 3,3 \rangle\} \\ A_3 &= \{\langle 1,3 \rangle, \langle 2,2 \rangle, \langle 3,1 \rangle\} \\ A_4 &= \{\langle 1,2 \rangle, \langle 2,3 \rangle, \langle 3,1 \rangle\} \\ A_5 &= \{\langle 1,1 \rangle, \langle 2,3 \rangle, \langle 3,2 \rangle\} \\ A_6 &= \{\langle 1,3 \rangle, \langle 2,1 \rangle, \langle 3,2 \rangle\}. \end{aligned} \tag{5}$$

There are twelve elements in the set of unobserved complementarities, and according to Lemma 1 we only need to know the values of the four of them in

$$\begin{aligned} B &= (c(1,2,1,2), c(1,2,1,3), c(1,3,1,2), c(1,3,1,3)) = \\ &= (- (e_{\langle 1,2 \rangle} + e_{\langle 2,1 \rangle}), e_{\langle 2,3 \rangle} - (e_{\langle 1,3 \rangle} + e_{\langle 2,1 \rangle}), e_{\langle 3,2 \rangle} - (e_{\langle 1,2 \rangle} + e_{\langle 3,1 \rangle}), - (e_{\langle 1,3 \rangle} + e_{\langle 3,1 \rangle})). \end{aligned}$$

To verify Lemma 1, we can construct the rest of the elements of the vector  $C$  from the sums and

differences of the four unobserved complementarities in  $B$ . Here we present one example:

$$\begin{aligned}
c(2, 3, 2, 3) &= e_{\langle 2,2 \rangle} + e_{\langle 3,3 \rangle} - (e_{\langle 2,3 \rangle} + e_{\langle 3,2 \rangle}) \\
&= - (e_{\langle 2,3 \rangle} + e_{\langle 3,2 \rangle}) \\
&= c(1, 2, 1, 2) - c(1, 2, 1, 3) - c(1, 3, 1, 2) + c(1, 3, 1, 3),
\end{aligned}$$

where the first equality uses the definition of an unobserved complementarity, the second equality uses the location normalization (2), and the third equality uses algebra. To verify Lemma 2, additional algebra shows that

$$\begin{pmatrix} \tilde{S}(A_1, E) \\ \tilde{S}(A_2, E) \\ \tilde{S}(A_3, E) \\ \tilde{S}(A_4, E) \\ \tilde{S}(A_5, E) \\ \tilde{S}(A_6, E) \end{pmatrix} = \begin{pmatrix} 0 \\ e_{\langle 1,2 \rangle} + e_{\langle 2,1 \rangle} \\ e_{\langle 1,3 \rangle} + e_{\langle 3,1 \rangle} \\ e_{\langle 1,2 \rangle} + e_{\langle 2,3 \rangle} + e_{\langle 3,1 \rangle} \\ e_{\langle 2,3 \rangle} + e_{\langle 3,2 \rangle} \\ e_{\langle 1,3 \rangle} + e_{\langle 2,1 \rangle} + e_{\langle 3,2 \rangle} \end{pmatrix} = \begin{pmatrix} 0 \\ -c(1, 2, 1, 2) \\ -c(1, 3, 1, 3) \\ c(1, 2, 2, 3) - c(1, 3, 1, 3) \\ -c(2, 3, 2, 3) \\ -c(1, 3, 1, 3) + c(2, 3, 1, 2) \end{pmatrix}. \quad (6)$$

One direction of Lemma 3 states that, given two realizations  $C_1$  and  $C_2$ , if  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all  $A$ , then  $C_1 = C_2$ . Algebra shows that

$$\begin{aligned}
c(1, 2, 1, 2) &= \tilde{S}(A_1, C) - \tilde{S}(A_2, C) \\
c(1, 2, 1, 3) &= \tilde{S}(A_5, C) - \tilde{S}(A_6, C) \\
c(1, 3, 1, 2) &= \tilde{S}(A_5, C) - \tilde{S}(A_4, C) \\
c(1, 3, 1, 3) &= \tilde{S}(A_1, C) - \tilde{S}(A_3, C).
\end{aligned}$$

Therefore, if all the elements in  $B$  are equal for  $\tilde{S}(A, C_1)$  and  $\tilde{S}(A, C_2)$ ,  $C_1$  must equal  $C_2$  by Lemma 1. The other direction of Lemma 3 states that if  $C_1 = C_2$ , then  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all  $A$ . This follows from (6), recalling that  $\tilde{S}(A, E)$  and  $\tilde{S}(A, C)$  are overloaded notation for the same sum of unobserved production.

## 2.5 Identifying the Distribution of Unobserved Complementarities

We wish to identify the CDF  $F(C)$  of unobserved complementarities. There are  $N!$  possible assignments in a one-to-one matching game without unmatched agents. A stable assignment  $A$  for a market  $(E, Z)$  maximizes  $S(A, E, Z)$ , so differences in assignment production  $S(A_2, E, Z) - S(A_1, E, Z)$  govern the pairwise stable assignment. We investigate the random variable that is the difference in the unobserved production of each assignment  $A$  and the unobserved production of  $A_1 = \{\langle 1, 1 \rangle, \dots, \langle N, N \rangle\}$ ,

where  $\tilde{S}(A_1, C) = 0$  for all  $C$  by the location normalization (2). Let  $\tilde{S} = (\tilde{S}(A_i, C))_{i=2}^{N!}$  be a vector of random variables giving the unobserved production of all assignments in a matching market. In order to identify  $F(C)$ , we first trace the CDF  $H(\tilde{S})$  by varying  $Z$  and then perform a change of variables from  $\tilde{S}$  to  $C$ .

**Lemma 4.** *The CDF  $H(\tilde{S})$  of unobserved production for all assignments is identified.*

*Proof.* Let a primitive unobservable  $E^*$  give the unobserved complementarities  $C^*$  and the corresponding vector of the unobserved production  $\tilde{S}^*$ . Set

$$z_{\langle u, d \rangle}^* = -e_{\langle u, d \rangle}^*.$$

Then by the location normalization,  $S(A_1, E, Z^*) = 0, \forall E$  at  $Z^*$ , and

$$S(A, E^*, Z^*) = \tilde{S}(A, C^*) + \sum_{\langle u, d \rangle \in A} z_{\langle u, d \rangle}^* = 0 \forall A.$$

Therefore for any  $A \neq A_1$ ,  $\tilde{S}(A, C) \leq \tilde{S}(A, C^*)$  for some  $C$  generated by some  $E$  if and only if

$$S(A, E, Z^*) \leq S(A, E^*, Z^*) = 0 = S(A_1, E, Z^*).$$

At  $\tilde{S}^*$ , by definition the joint CDF (using the one-to-one change of variables between  $C$  and  $S$ ) is

$$H(\tilde{S}^*) = \Pr(\tilde{S}(A, C) \leq \tilde{S}(A, C^*), \forall A \neq A_1).$$

Therefore

$$\begin{aligned} H(\tilde{S}^*) &= \Pr(\tilde{S}(A, C) \leq \tilde{S}(A, C^*), \forall A \neq A_1) \\ &= \Pr(S(A, E, Z^*) \leq S(A, E^*, Z^*), \forall A \neq A_1 \mid Z^*) \\ &= \Pr(S(A, E, Z^*) \leq 0, \forall A \neq A_1 \mid Z^*) \\ &= \Pr(S(A, E, Z^*) \leq S(A_1, E, Z^*), \forall A \neq A_1 \mid Z^*) \\ &= \Pr(A_1 \mid Z^*). \end{aligned}$$

The second equality adds the production from  $Z^*$  for each assignment  $A$  to both sides of each inequality; the third equality uses the particular choice  $Z^*$  and the properties involving it derived just above; the fourth inequality uses the location normalization (2) that the sum of unobserved production for  $A_1$  for any  $E$  is 0 and that consequently  $z_{\langle u, d \rangle}^*$  for matches  $\langle u, d \rangle \in A_1$  are also set to 0 by the choice  $z_{\langle u, d \rangle}^* = -e_{\langle u, d \rangle}^*$  above; the fifth inequality is the definition of assignment  $A_1$  occurring given observables  $Z^*$ . Values of  $\tilde{S}^*$  that cannot be formed from a valid  $E^*$  will not occur.  $\square$

The proof of this lemma shows that identification of  $H(\tilde{S})$  uses data on the moments  $\Pr(A_1 | Z)$  for many values  $Z$ , where assignment  $A_1$  is the assignment whose unobserved production  $\tilde{S}(A_1, E) = 0 \forall E$  in the location normalization (2). The identification of  $H(\tilde{S})$  does not use the probabilities of other assignments. This is analogous to results on the multinomial choice model, where data on the probability of only one choice is needed for identification (e.g., Thompson, 1989).

It is crucial to have large and product support on  $Z$  to point identify  $H(\tilde{S})$  without restrictions on the support of  $E$ , again as in the binary and multinomial choice literature (Ichimura and Thompson, 1998; Lewbel, 1998, 2000; Matzkin, 2007; Gautier and Kitamura, 2011; Berry and Haile, 2010; Fox and Gandhi, 2012). For each point  $\tilde{S}$  for which we wish to identify the value  $H(\tilde{S})$ , the proof of Lemma 4 produces a corresponding  $Z$  and hence a  $\Pr(A_1 | Z)$ . If that value  $Z$  is not in the support of the data, the identification strategy in the proof will not point identify  $H(\tilde{S})$  for this particular  $\tilde{S}$ .

**Theorem 2.** *The distribution  $F(C)$  of market-level unobserved complementarities is identified in a matching game where all agents must be matched.*

*Proof.* By Lemma 3, there exists a one-to-one and onto correspondence  $J$  from the space of  $C$  to the space of  $\tilde{S}$ . By Lemma 4, the distribution of  $\tilde{S}$ ,  $H(\tilde{S})$  is identified. We can perform a change of variables and learn the distribution of  $C$ : every valid  $C$  corresponds with a  $\tilde{S} = J(C)$ , and  $F(C) = H(J(C))$  can be traced. Values of  $C$  that cannot be reconciled with a valid  $E$  and a valid  $\tilde{S}$  will not occur.  $\square$

Our results imply that only one  $F(C)$  can generate the limiting information on  $\Pr(A | Z)$  for pairs  $(A, Z)$ . As discussed previously, various nonparametric mixtures estimators can be used with a finite sample of data. Again, we recommend estimating  $F_B(B)$  and not  $F(C)$ .

Our identification results are for the general case where agent indices such as  $u$  and  $d$  have common meanings across markets. For example, in Fox (2010a) a matching market is an automotive supplier component category and the same suppliers operate in multiple component categories. In applications where agent indices are arbitrary because the same agents do not operate in different markets, one should additionally impose that  $F(C)$  is exchangeable in the agent indices. An exchangeable distribution  $F(C)$  is a special case of our identification results.<sup>12</sup>

### 3 Extensions

We consider various extensions to the one-to-one matching game where all agents are matched.

<sup>12</sup>A matching market where agent indices have no meaning is analogous to a multinomial choice problem when each agent is offered a set of choices where the indices of choices have no meaning. Even if indices are arbitrary, one should not relabel indices after observing values in  $Z$  or in  $A$ , for example by setting the firm with the highest value for a  $z_{\langle u, d \rangle}$  to always be firm 1. In this case, the support of  $Z$  will not be  $\mathbb{R}^{N^2}$ .



### 3.1 Other Observed Variables $X$ and Random Preferences

In addition to the match-specific special regressors  $Z$ , researchers often observe other match-specific and agent-specific data, which we collect in the random variable  $X$ . We also include in  $X$  the number of agents on each side,  $N$ , to allow the size of the market to vary across the sample. An example of a production function augmented by the elements of  $X$  is

$$(x_u \cdot x_d)' \beta_{\langle u,d \rangle,1} + x'_{\langle u,d \rangle} \beta_{\langle u,d \rangle,2} + \mu_{\langle u,d \rangle} + z_{\langle u,d \rangle},$$

where  $x_u$  is a vector of upstream firm characteristics,  $x_d$  is a vector of downstream firm characteristics,  $x_u \cdot x_d$  is a vector of all interactions between upstream and downstream characteristics,  $x_{\langle u,d \rangle}$  is a vector of match-specific characteristics,  $\mu_{\langle u,d \rangle}$  is a random intercept capturing unobserved characteristics of both  $u$  and  $d$ , and  $\beta_{\langle u,d \rangle,1}$  and  $\beta_{\langle u,d \rangle,2}$  are random coefficient vectors specific to the match. The two random coefficient vectors can be the sum of the **random preferences** of upstream and downstream firms for own and partner characteristics. In a marriage setting, we allow men to have heterogeneous preferences over the observed characteristics of women, which is one explanation for why observationally identical men marry observationally distinct women.

In this example,

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

Now we define

$$e_{\langle u,d \rangle} = (x_u \cdot x_d)' \beta_{\langle u,d \rangle,1} + x'_{\langle u,d \rangle} \beta_{\langle u,d \rangle,2} + \mu_{\langle u,d \rangle}$$

and

$$c(u_1, u_2, d_1, d_2) = e_{\langle u_1, d_1 \rangle} + e_{\langle u_2, d_2 \rangle} - e_{\langle u_1, d_2 \rangle} - e_{\langle u_2, d_1 \rangle}.$$

Using the same notation as before, we define the vector of unobserved complementarities as (4). This definition of  $C$  now depends on the realizations of  $X$ . Our previous argument in Theorem 2 does not use  $X$ , therefore we can condition on a realization of  $X$  to identify the conditional-on- $X$  distribution of unobserved complementarities  $F(C | X)$ . We of course require variation in  $Z$  as before, but now  $Z$  must have full support conditional on each realization of  $X$ . We do not require that  $C$  and  $E$  are independent of  $X$ , but both unobservables must still be independent of  $Z$  conditional on  $X$ .

**Corollary 1.** *The distribution  $F(C | X)$  of market-level unobserved complementarities conditional on  $X$  is identified in a one-to-one matching game where all agents must be matched.*

Our identification of distributions of heterogeneity conditional on  $X$  follows arguments in the multinomial choice literature (Lewbel, 2000; Matzkin, 2007; Berry and Haile, 2010). If instead we assumed that  $\beta_{\langle u,d \rangle,1}$ ,  $\beta_{\langle u,d \rangle,2}$  and  $\mu_{\langle u,d \rangle}$  are independent of  $X$  and tried to identify the distribution of  $C(X)$  as a vector of functions of  $X$ , we could use the identification framework of Fox and Gandhi (2012). A previous version of our paper used stronger assumptions to identify distributions of functions

of  $X$ .

### 3.2 Inclusion of Heterogeneous Coefficients on $z$

We define the production to a match  $\langle u, d \rangle$  to be

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

where  $\gamma_{\langle u, d \rangle} \neq 0$  is a match-specific coefficient. The coefficients  $\gamma_{\langle u, d \rangle}$  vary across matches within each market but not across markets. Therefore, the  $\gamma_{\langle u, d \rangle}$  are fixed parameters to be identified and not random coefficients. Fixing coefficients across markets but not within markets only makes sense in a context where firm indices like  $u$  and  $d$  have a consistent meaning across markets. For example, the same set of upstream and downstream firms may participate in multiple matching markets, as in Fox (2010a), where each market is a separate automotive component category.<sup>13</sup>

We apply a scale normalization on production by setting  $|\gamma_{\langle 1, 1 \rangle}| = 1$ . Because of transferable utility, we can identify the relative scale of each match's production. We use the matrix  $\Gamma = (\gamma_{\langle u, d \rangle})_{u, d \in N}$ . The terms  $C$  and  $X$  are defined as before.

**Assumption 1.** *Given any  $\delta$ ,  $0 < \delta < 1$ , there exists  $y_1, y_2$  such that for any  $A_1, A_2$ , and  $X$ , and for all  $a > y_1$ , and  $b < y_2$ ,*

$$\Pr(\tilde{S}(A_1, C) - \tilde{S}(A_2, C) < a \mid X) > \delta$$

and

$$\Pr(\tilde{S}(A_1, C) - \tilde{S}(A_2, C) < b \mid X) < \delta.$$

This assumption rules out probability masses at the infinities of  $\tilde{S}(A_1, C) - \tilde{S}(A_2, C)$ . Under this assumption, we obtain identification of  $\Gamma$  and  $F(C \mid X)$ .

#### Theorem 3.

1. *The parameter matrix  $\Gamma$  is identified in a one-to-one matching game where all agents must be matched.*
2. *Assume that no element of  $\Gamma$  is 0. In the same setting, the distribution  $F(C \mid X)$  of market-level unobserved complementarities conditional on  $X$  is identified once  $\Gamma$  is identified.*

---

<sup>13</sup>In a marriage setting with different individuals in each market, we could assume that  $\gamma_{\langle u, d \rangle}$  is the same for all matches where the men are all in the same demographic class (such as college graduates) and the women are all in the same demographic class (such as high-school graduates). This suggested use of demographic classes is partially reminiscent of Chiappori et al. (2010b), who use data over time on the US marriage market to estimate a different variance of the type I extreme value (logit) utility errors in a Choo and Siow (2006) style model for each male demographic class and for each female demographic class. Our suggested approach in this footnote lets  $\gamma_{\langle u, d \rangle}$  vary by the intersection of male and female demographic classes.

The proof is included in the appendix. While Theorem 3 identifies a more general model than Theorem 2, the identification proof for the matrix  $\Gamma$  in Theorem 3 operates by contradiction. Once  $\Gamma$  is identified, a variant of the proof of Theorem 2 can be used to trace  $F(C | X)$ . We rule out that any  $\gamma_{\langle u, d \rangle}$  is 0 in order to use the special regressors in  $Z$  to trace  $F(C | X)$ .

### 3.3 Agent-Specific Characteristics

Match specific  $z$ 's with full support are not always available in datasets. Instead, we now assume the researcher has scalar agent-specific regressors  $z_u$  and  $z_d$ , which enter the production function

$$e_u \cdot e_d + z_u \cdot z_d, \quad (7)$$

where  $e_u$  and  $e_d$  are unobserved agent-specific characteristics. We still require that the  $z$ 's have full and product support, meaning

$$Z = ((z_u)_{u \in N}, (z_d)_{d \in N})$$

now has support on  $\mathbb{R}^{2N}$ .

We need location and scale normalizations. We first normalize  $e_u = 0$  for  $u = 1$  by subtracting an equal amount from all  $e_u$ . This location normalization does not affect the optimal assignment: the sum of unobserved production for a given assignment  $A$  is decreased by a constant times the sum of the  $e_d$ , which is independent of  $A$ . Therefore, the ordering of the total unobserved production for the different assignments does not change. Similarly, we normalize  $e_d = 0$  for  $d = 1$ . Next, we apply a scale normalization and set  $e_u = 1$  for  $u = 2$ . This is equivalent to setting  $\tilde{e}_u = \frac{e_u}{e_{u=2}}$  and  $\tilde{e}_d = e_d \cdot e_{u=2}$ . Because  $F(C|X)$  has continuous support, the probability of  $e_{u=2} = 0$  is 0. This scale normalization keeps the production for each match the same, as  $\tilde{e}_u \cdot \tilde{e}_d = e_u \cdot e_d$ . Given the normalizations, we, for this subsection only, redefine a market-level type to be

$$E = ((e_u)_{u=3}^N, (e_d)_{d=2}^N). \quad (8)$$

In other words,  $E$  is comprised of  $2N - 3$  unobservables.  $X$  is composed of agent-specific observable characteristics entering the indices  $e_u$  and  $e_d$ . For example,  $e_u = x'_u \beta_u + \mu_u$ , where the vector  $x_u$  is comprised of observable characteristics of firm  $u$  other than  $z_u$ ,  $\beta_u$  is a vector of possibly random coefficients, and  $\mu_u$  is a scalar random intercept. We show that the distribution of  $E$  is identified.

**Theorem 4.** *The distribution  $G(E | X)$  is identified in the one-to-one matching model with agent-specific characteristics, agent-specific unobservables, and without unmatched agents.*

The proof is in the appendix. We emphasize that we identify the distribution of agent-specific characteristics and not just the distribution of unobserved complementarities. The identification of the distribution of unobserved complementarities  $F(C | X)$  follows from the identification of  $G(E | X)$ .

The scale normalization for  $E$  is with respect to the second upstream firm. The scale of the distribution  $G(E | X)$  has a more meaningful interpretation when the second upstream firm is the same firm across matching markets. If the firm indexed as the second firm in each matching market is arbitrary, all we learn is the distribution of unobserved characteristics relative to some arbitrary firm in each market. This makes comparisons across markets difficult.<sup>14</sup> As before, we should impose exchangeability of  $G(E | X)$  in the case where agent indices have no underlying meaning. We can always use the identified  $G(E | X)$  to compute counterfactual assignment probabilities.

## 4 Data on Unmatched Agents

Return to the case of match-specific unobservables and match-specific special regressors. Up until this point, we consider matching games where all agents have to be matched. This assumption is reasonable when only data on observed matches are available. For example, it may be unreasonable to assume that data on all potential entrants to a matching market exist. In some situations, however, researchers can also observe the identities of unmatched agents. Data are available, for example, on potential merger partners or on single people in a marriage market. When data on unmatched agents do exist, we can go beyond unobserved complementarities and identify the distribution of match-specific unobservables.

Here,  $X$  can contain separate numbers of downstream firms  $N_d$  and upstream firms  $N_u$ . Use  $\langle u, 0 \rangle$  and  $\langle 0, d \rangle$  to denote an upstream firm and a downstream firm that are not matched. An assignment  $A$  can be  $\{\langle u_1, 0 \rangle, \langle u_2, d_2 \rangle, \langle 0, d_2 \rangle\}$ , allowing single firms. We normalize the production of unmatched agents to be 0, instead of the location normalization in (2). Therefore,

$$E = \begin{pmatrix} e_{\langle 1, 1 \rangle} & \cdots & e_{\langle 1, N_d \rangle} \\ \vdots & \ddots & \vdots \\ e_{\langle N_u, 1 \rangle} & \cdots & e_{\langle N_u, N_d \rangle} \end{pmatrix}.$$

We do not need match-specific regressors  $z_{\langle u, 0 \rangle}$  and  $z_{\langle 0, d \rangle}$  for unmatched firms; they can be included in  $X$  if present.

The data generating process is still (1). One difference is that a pairwise stable assignment needs to satisfy individual rationality: each non-singleton realized match has production greater than 0.

**Theorem 5.** *The distribution  $G(E | X)$  of market-level unobservables is constructively identified with data on unmatched agents.*

*Proof.* Condition on  $X$ . Let  $A_0$  denote the assignment where no agents are matched, then  $\tilde{S}(A_0, E) =$

---

<sup>14</sup>An analogous issue happens in discrete choice models. If one normalizes the utility of each choice relative to choice 1 and choice 1 has no common interpretation across agents, the joint (across choices) distribution of utility is hard to compare across agents.

0 for all  $E$ . Let  $E^*$  be an arbitrary realization. Let  $Z^* = \left( z_{\langle u, d \rangle}^* \right)_{u, d \in N}$  be such that  $z_{\langle u, d \rangle}^* = -e_{\langle u, d \rangle}^*$ . Then  $S(A, Z^*, E^*) = 0$  for all  $A$ , and  $S(A_0, Z^*, E) = 0$  for any  $E$ . Thus for all  $A$  and all  $E \leq E^*$  elementwise,  $S(A, Z^*, E) \leq 0 = S(A_0, Z^*, E)$ . Therefore  $G(E^*) = \Pr(A_0 | Z^*)$ .  $\square$

The proof shows that the distribution  $G(E | X)$  for some  $X$  can be traced using the probability that all agents are single, conditional on  $Z$ . The individual rationality decision to be single identifies  $G(E | X)$  while the sorting of matched firms to other matched firms identifies only  $F(C | X)$ . Using an individual rationality condition is more similar to the utility maximization assumptions used in the identification of single-agent discrete choice models and discrete Nash games (Lewbel, 2000; Matzkin, 2007; Berry and Haile, 2010; Berry and Tamer, 2006).

## 5 One-Sided and Many-to-Many Matching

One-sided matching is more general than two-sided matching. Agents are not divided into men and women or upstream and downstream firms; all matches with at most two agents can hypothetically occur. Examples include models with homosexual relationships and models of mergers, where which firm is a target and which firm is an acquirer are not defined as fixed roles. In one-sided matching, a pairwise stable assignment may not exist (Chiappori, Galichon and Salanié, 2012). However, if an assignment does exist in a one-to-one matching game, it is unique with probability 1 by the same social planner arguments we cite for two-sided matching. None of our identification arguments use the two sides of the market in a fundamental way. Therefore, all our arguments generalize to one-sided matching when a pairwise stable assignment exists. Under non-existence, probabilities will not sum to 1. Relatedly, Ciliberto and Tamer (2009) estimate a discrete Nash game of perfect information assuming pure strategies, while proving existence of a Nash equilibrium requires mixed strategies.

Many-to-many matching where the production of each match does not depend on other matches (the underlying profit of each firm is additively separable across matches and hence allows the definition of production functions at the match level) is a simple generalization of the matching models studied in this paper. Sotomayor (1999) shows that existence of a pairwise stable assignment is guaranteed and similar arguments to ours show it is unique with probability 1. All of our lemmas and theorems can be shown to extend to this case, where the production of each match does not depend on other matches. In particular, say the number of matches each agent makes in a feasible assignment is not always equal to the maximum number of matches that the agent can physically make. Then there is an element of individual rationality in the decision not to form all possible matches, and an argument similar to that in the proof of Theorem 5 can be used.

Many-to-many matching where the production of each match depends on other matches involving the same agents is more complex. Under a theoretical restriction known as *substitutes* (which rules out same-side complementarities), a pairwise stable assignment will exist and be unique with prob-

ability 1, as all pairwise stable assignments are fully stable and hence maximize the social planner’s objective function (Hatfield and Milgrom, 2005; Hatfield et al., 2011). Consider an example where each upstream firm  $u$  can match with two downstream firms  $d_1$  and  $d_2$ . If we impose the substitutes condition, we can work with the production  $z_{\langle u, d_1, d_2 \rangle} + e_{\langle u, d_1, d_2 \rangle}$  and identify the distribution of the appropriate definition of unobserved complementarities or the distribution of  $e_{\langle u, d_1, d_2 \rangle}$ , depending on whether an identification argument based on individual rationality can be used. We need one special regressor  $z$  for each  $e$  term; the dimension of the data must be analogous to the dimension of the unobservables.

If we do not impose the substitutes condition, the many-to-many matching problem will have cases of non-existence and will also have multiple pairwise stable assignments with positive probability. Methods used in the literature on estimating Nash games with multiple equilibria will likely need to be employed, which is beyond the scope of this paper (Bajari et al., 2010; Beresteanu et al., 2011; Ciliberto and Tamer, 2009; Galichon and Henry, 2011). Another approach to multiplicity is to impose a particular model of search and match formation, as in Christakis et al. (2010). Also, the maximum score approach of Fox (2010b) uses broad assumptions on equilibrium selection to gain identification without ruling out multiplicity.

## 6 Conclusion

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

Using data on only matched firms, one can identify distributions of what we call unobserved complementarities but not the underlying primitive distribution of match-specific unobservables. The distribution of complementarities is enough to compute assignment production levels and therefore counterfactual assignment probabilities.

In extensions, we can include other covariates  $X$  and identify distributions of unobservable complementarities conditional on  $X$ . We show that it is possible to identify heterogeneous-within-a-market coefficients on the special regressors. We also examine a model where firm-level observed and unobserved characteristics both enter the production function multiplicatively, and we show identification of the distribution of the firm level unobservables.

Finally, if the data contain unmatched firms, the individual rationality decision to not be unmatched helps identify the distribution of primitively specified unobservables, not just the distribution of unobserved complementarities.

## A Appendix

### A.1 Proof of Lemma 1

We construct a vector of unobserved complementarities such that any unobserved complementarity is equal to sums and differences of elements in this vector.

The vector we construct is

$$\begin{aligned} B &= (c(u_1, u_2, d_1, d_2) \mid u_1 = d_1 = 1, \forall u_2, d_2 \in \{2, \dots, N\}) \\ &= (c(1, u_2, 1, d_2)). \end{aligned}$$

There are  $(N - 1)^2$  elements in  $B$ .

First we show that any unobserved complementarity of the form  $c(1, u, d_1, d_2)$  is equal to the difference of two elements of  $B$ :

$$\begin{aligned} c(1, u, d_1, d_2) &= e_{\langle 1, d_1 \rangle} + e_{\langle u, d_2 \rangle} - (e_{\langle 1, d_2 \rangle} + e_{\langle u, d_1 \rangle}) \\ &= e_{\langle 1, 1 \rangle} + e_{\langle u, d_2 \rangle} - (e_{\langle 1, d_2 \rangle} + e_{\langle u, 1 \rangle}) - (e_{\langle 1, 1 \rangle} + e_{\langle u, d_1 \rangle} - (e_{\langle 1, d_1 \rangle} + e_{\langle u, 1 \rangle})) \\ &= c(1, u, 1, d_2) - c(1, u, 1, d_1). \end{aligned}$$

Next,

$$\begin{aligned} c(u_1, u_2, d_1, d_2) &= e_{\langle u_1, d_1 \rangle} + e_{\langle u_2, d_2 \rangle} - (e_{\langle u_1, d_2 \rangle} + e_{\langle u_2, d_1 \rangle}) \\ &= e_{\langle 1, d_1 \rangle} + e_{\langle u_2, d_2 \rangle} - (e_{\langle 1, d_2 \rangle} + e_{\langle u_2, d_1 \rangle}) - (e_{\langle 1, d_1 \rangle} + e_{\langle u_1, d_2 \rangle} - (e_{\langle 1, d_2 \rangle} + e_{\langle u_1, d_1 \rangle})) \\ &= c(1, u_2, d_1, d_2) - c(1, u_1, d_1, d_2). \end{aligned}$$

Because we have shown that any  $c(1, u_2, d_1, d_2)$  is a difference of two elements in  $B$ ,  $c(u_1, u_2, d_1, d_2)$  can be written as the sums and differences of elements in  $B$ .

### A.2 Proof of Lemma 2

We prove the lemma by mathematical induction on  $N$ , the number of upstream firms, which is equal to the number of downstream firms. When  $N = 2$ , in assignment  $A_1 = \{\langle 1, 1 \rangle, \langle 2, 2 \rangle\}$   $\tilde{S}(A_1, E)$  is 0 by our location normalization, and in  $A_2 = \{\langle 1, 2 \rangle, \langle 2, 1 \rangle\}$   $\tilde{S}(A_2, E) = e_{\langle 1, 2 \rangle} + e_{\langle 2, 1 \rangle} = -c(1, 2, 1, 2)$ . Thus the lemma is true for  $N = 2$ . The induction hypothesis is that the lemma is true for  $N - 1$  for some arbitrary  $N$ . For the case of markets with  $N$  upstream firms, there are two different types of assignments  $A$  that can occur. We use the notation of  $A(u) = d$  if match  $\langle u, d \rangle$  occurs in  $A$ .

*Case 1.*  $A$  is such that  $A(N) = N$ , where again  $N$  is the last upstream firm (and hence downstream

firm) index. Then

$$\tilde{S}(A, E) = \sum_{u=1}^{N-1} e_{\langle u, A(u) \rangle} + e_{\langle N, N \rangle} = \sum_{u=1}^{N-1} e_{\langle u, A(u) \rangle} + 0,$$

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

$$\tilde{S}(A, E) = \sum_{u=1}^{N-1} e_{\langle u, A(u) \rangle} + 0 = \sum_{u=1}^{N-1} e_{\langle u, A^*(u) \rangle}.$$

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

*Case 2.*  $A$  is such that  $A(N) = d^*$ , where  $d^* < N$ . Because all agents are matched, among upstream firms  $1, \dots, N-1$  there exists an upstream firm  $u^*$  who matches with the last downstream firm  $N$ :  $A(u^*) = N$ . Therefore

$$\begin{aligned} \tilde{S}(A, E) &= \sum_{\substack{u=1, \\ u \neq u^*}}^{N-1} e_{\langle u, A(u) \rangle} + e_{\langle N, A(N) \rangle} + e_{\langle u^*, A(u^*) \rangle} \\ &= \sum_{\substack{u=1, \\ u \neq u^*}}^{N-1} e_{\langle u, A(u) \rangle} + e_{\langle N, d^* \rangle} + e_{\langle u^*, N \rangle}. \end{aligned} \tag{9}$$

Recall  $c(N, u^*, d^*, N) = e_{\langle N, d^* \rangle} + e_{\langle u^*, N \rangle} - (e_{\langle N, N \rangle} + e_{\langle u^*, d^* \rangle})$ . Because  $e_{\langle N, N \rangle} = 0$  by the location normalization, we have

$$e_{\langle N, d^* \rangle} + e_{\langle u^*, N \rangle} = c(N, u^*, d^*, N) + e_{\langle u^*, d^* \rangle}.$$

Substitute this equation into (9):

$$\tilde{S}(A, E) = \sum_{\substack{u=1, \\ u \neq u^*}}^{N-1} e_{\langle u, A(u) \rangle} + e_{\langle u^*, d^* \rangle} + c(N, u^*, d^*, N).$$



Notice that the first  $(N - 1)$  terms only involve the first  $(N - 1)$  upstream firms and the first  $(N - 1)$  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  $(N - 1)$ -by- $(N - 1)$  submarket. By the induction hypothesis, this is a sum of unobserved complementarities. Therefore  $\tilde{S}(A, E)$  is a sum of unobserved complementarities.

The induction is complete, and the lemma is true for all  $N$ .

### A.3 Proof of Lemma 3

In proving the result, we use the same notation  $A(u) = d$  as in the previous proof. The “only if” direction is a result of Lemma 2:  $\tilde{S}(A, C)$  can be expressed as a sum of  $c(\cdot)$  that is not specific to the realization  $E$  (or  $C$ ). Suppose  $C_1 = C_2$ . Because  $c^1(u_1, u_2, d_1, d_2) = c^2(u_1, u_2, d_1, d_2)$  for all sets of two upstream and two downstream firms, then  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all assignments  $A$ . To prove the “if” direction, suppose  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all  $A$ . Given any upstream firms  $u_1, u_2$  and downstream firms  $d_1, d_2$ , we want to show that  $c^1(u_1, u_2, d_1, d_2) = c^2(u_1, u_2, d_1, d_2)$ , where the superscripts index  $C_1$  and  $C_2$ . Consider assignment  $A_1$  such that  $A_1(u_1) = d_1$  and  $A_1(u_2) = d_2$  and the assignment  $A_2$  such that  $A_2(u_1) = d_2$  and  $A_2(u_2) = d_1$ , where  $A_1(u) = A_2(u)$  for all other upstream firms  $u$ . As  $\tilde{S}(A, C_1) = \tilde{S}(A, C_2)$  for all  $A$ ,

$$\begin{aligned} e_{\langle u_1, A_1(u_1) \rangle}^1 + e_{\langle u_2, A_1(u_2) \rangle}^1 + \sum_{u \neq u_1, u_2} e_{\langle u, A_1(u) \rangle}^1 &= e_{\langle u_1, A_1(u_1) \rangle}^2 + e_{\langle u_2, A_1(u_2) \rangle}^2 + \sum_{u \neq u_1, u_2} e_{\langle u, A_1(u) \rangle}^2 \\ e_{\langle u_1, A_2(u_1) \rangle}^1 + e_{\langle u_2, A_2(u_2) \rangle}^1 + \sum_{u \neq u_1, u_2} e_{\langle u, A_2(u) \rangle}^1 &= e_{\langle u_1, A_2(u_1) \rangle}^2 + e_{\langle u_2, A_2(u_2) \rangle}^2 + \sum_{u \neq u_1, u_2} e_{\langle u, A_2(u) \rangle}^2. \end{aligned}$$

Subtracting the equations will cancel out all the terms in the summations. Substitute in the values for  $A_1(\cdot)$  and  $A_2(\cdot)$  to give

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

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

### A.4 Proof of Theorem 3

We first state a lemma that is helpful in the proof of the theorem. In the lemma,  $A_0$  is an arbitrary assignment, not necessarily the assignment  $\{\langle 1, 1 \rangle, \dots, \langle N, N \rangle\}$ .

**Lemma 5.** *Let  $0 < \delta < 1$  be given. Given Assumption 1 and a particular, arbitrary assignment  $A_0$ ,  $\exists \zeta_1, \zeta_2$ , such that for any  $X$ , and for all  $a > \zeta_1$ , and  $b < \zeta_2$ ,*

$$\Pr(\tilde{S}(A, E) - \tilde{S}(A_0, E) < a, \forall A \neq A_0 \mid X) > \delta$$

and

$$\Pr(\tilde{S}(A, E) - \tilde{S}(A_0, E) < b, \forall A \neq A_0 \mid X) < \delta.$$

*Proof.* We construct  $\zeta_1$  inductively. List all possible assignments as  $A_0, A_1, \dots, A_{N!-1}$ . Denote the event  $\tilde{S}(A_i, \theta) - \tilde{S}(A_0, \theta) < a$  as  $B_i(a)$ . Denote  $C_n(a) = \cap_{i=1 \dots n} B_i(a)$ . The first part of the lemma is to show that there exists  $\zeta_1$  such that  $\forall a > \zeta_1$ ,  $\Pr(C_{N!-1}(a) \mid X) > \delta$ . Given  $\delta$ , let  $\delta_1 = \frac{1}{2}(1 + \delta)$ . By Assumption 1, there exists  $y_1$  and  $y_2$  such that for  $a > y_1$   $\Pr(B_1(a) \mid X) > \delta_1$  and for  $a > y_2$   $\Pr(B_2(a) \mid X) > \delta_1$ . Let  $\zeta_1^{(2)} = \max(y_1, y_2)$ , then for any  $a > \zeta_1^{(2)}$

$$\begin{aligned} \Pr(C_2(a) \mid X) &= \Pr(B_1(a) \mid X) + \Pr(B_2(a) \mid X) - \Pr(B_1(a) \cup B_2(a) \mid X) \\ &> \Pr(B_1(a) \mid X) + \Pr(B_2(a) \mid X) - 1 \\ &\geq 2\delta_1 - 1 = \delta. \end{aligned}$$

Given  $\delta$  and  $n-1$ , then by induction there exists  $\zeta_1^{(n-1)}$  such that  $\forall a > \zeta_1^{(n-1)}$ ,  $\Pr(C_{n-1}(a) \mid X) > \delta$ . We do not use this hypothesis for this value of  $\delta$ . Instead, because the lemma applies to any value of  $\delta$ , we use the induction hypothesis for the value  $\delta_1$  instead. Therefore, there exists  $y_1^{(n)}$  such that for  $a > y_1^{(n)}$ ,  $\Pr(C_{n-1}(a) \mid X) > \delta_1$ . By Assumption 1, there exists  $y_2^{(n)}$  such that for  $a > y_2^{(n)}$ ,  $\Pr(B_n(a) \mid X) > \delta_1$ . Let  $\zeta_1^{(n)} = \max(y_1^{(n)}, y_2^{(n)})$ , then just as above,  $\Pr(C_n(a) \mid X) > \delta$ . Therefore the induction is complete; denote  $\zeta_1 = \zeta_1^{(N!-1)}$ . For  $a > \zeta_1$ ,  $\Pr(C_{N!-1}(a) \mid X) > \delta$ .

Any  $y_2$  satisfying the second part of Assumption 1 is  $\zeta_2$ : for given  $A$  and  $\delta$ , there exists  $y_2$  such that  $\forall b < y_2$ ,  $\Pr(\tilde{S}(A, \psi) - \tilde{S}(A_0, \psi) < b \mid X) < \delta$ . The intersection of these sets over  $A$  has a weakly smaller probability than  $\delta$ .  $\square$

Now we are ready to state the proof of Theorem 3.

*Proof.* Condition on  $X$  if it is present. The location normalization is that  $|\gamma_{\langle 1, 1 \rangle}| = 1$ . We first prove the theorem with  $\gamma_{\langle 1, 1 \rangle} = 1$ . The argument is similar with  $\gamma_{\langle 1, 1 \rangle} = -1$ . Consider two parameter pairs  $(\Gamma, F)$  and  $(\hat{\Gamma}, \hat{F})$ . If  $\Gamma = \hat{\Gamma}$ , a variant of the proof of Theorem 2 can be used to trace out  $F$  if each element of  $\Gamma$  is nonzero, thus uniquely identifying it. So the interesting case is when  $\Gamma \neq \hat{\Gamma}$ .

The majority of this proof shows the identification of  $\gamma_{\langle u_1, d_1 \rangle}$ , where either  $u_1 = 1$  or  $d_1 = 1$ . Let  $\Pr(A \mid Z)$  represent probabilities of an assignment (integrating over  $E$ ) under  $(\Gamma, F)$  and let  $\hat{\Pr}(A \mid Z)$  represent probabilities under  $(\hat{\Gamma}, \hat{F})$ . We will find an assignment  $A^*$  and data matrix  $\check{Z}$  where  $\Pr(A^* \mid \check{Z}) \neq \hat{\Pr}(A^* \mid \check{Z})$ , thus proving identification. The identification of parameters other than  $\gamma_{\langle u_1, d_1 \rangle}$  will follow similar arguments, as we discuss at the end.

Therefore, suppose  $\gamma_{\langle u_1, d_1 \rangle} \neq \hat{\gamma}_{\langle u_1, d_1 \rangle}$ . Pick an assignment  $A^*$  that contains the match  $\langle 1, 1 \rangle$ . Then  $\langle u_1, d_1 \rangle$  is not a match of  $A^*$ , as each upstream firm and each downstream firm can be matched only once. Fix a positive number  $\delta$ ,  $0 < \delta < 1$ . Pick  $Z^*$  as follows:

1.  $z_{\langle u, d \rangle}^* = 0$  if  $\langle u, d \rangle \neq \langle u_1, d_1 \rangle$  or  $\langle u, d \rangle \notin A^*$ .
2. Without loss of generality, choose  $z_{\langle 1, 1 \rangle}^* > 0$  and  $z_{\langle u_1, d_1 \rangle}^*$  such that

$$\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^* - z_{\langle 1, 1 \rangle}^* > 0 > \hat{\gamma}_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^* - z_{\langle 1, 1 \rangle}^* \quad (10)$$

3. For any other  $\langle u, d \rangle \in A^*$ , choose  $z_{\langle u, d \rangle}^*$  such that  $\gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* > z_0 + |\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*|$ , where  $z_0$  is defined as follows. By Assumption 1 and Lemma 5, there exists  $z_0$  such that for any  $a \geq z_0$ ,  $\Pr(\tilde{S}(A, E) - \tilde{S}(A^*, E) < a, \forall A \neq A^* \mid X) > \delta$ .

Let the set

$$\mathcal{E}_{A^*} = \{E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) < z_0, \forall A \neq A^*\}.$$

Then  $\Pr(\mathcal{E}_{A^*}) > \delta$ . Denote the set of matches that occur in  $A^* \setminus \{\langle 1, 1 \rangle\}$  but not in  $A$  as  $A^* \setminus \{A \cup \langle 1, 1 \rangle\}$ . Any non- $A^*$  assignment  $A$  is in one of the following three mutually exclusive categories.

1.  $A$  contains the match  $\langle 1, 1 \rangle$ . Then  $A$  cannot contain the match  $\langle u_1, d_1 \rangle$ . Denote the set of such  $A$ 's as  $\mathcal{A}_1$ . Denote the set of  $E$ 's such that  $A^*$  has higher production than any  $A \in \mathcal{A}_1$  as

$$\mathcal{E}_1(Z^*) = \{E \mid S(A, E, Z^*) \leq S(A^*, E, Z^*), \forall A \in \mathcal{A}_1\},$$

which is equivalent to

$$\{E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq z_1^*(A), \forall A \in \mathcal{A}_1\},$$

where  $z_1^*(A) = \left( \sum_{\langle u, d \rangle \in A^* \setminus \{A \cup \langle 1, 1 \rangle\}} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right)$ . Because there are at least two different matches in the assignments  $A$  and  $A^*$ , there is at least one match in the set of  $A^* \setminus \{A \cup \langle 1, 1 \rangle\}$ . For each  $A \in \mathcal{A}_1$ , denote one such match as  $\langle \tilde{u}, \tilde{d} \rangle^A$  (keep in mind  $\langle \tilde{u}, \tilde{d} \rangle^A \in A^*$ , not  $A$ ). By Condition 3 in picking  $Z^*$ , for each  $A \in \mathcal{A}_1$

$$\begin{aligned} z_1^*(A) &> \gamma_{\langle \tilde{u}, \tilde{d} \rangle^A} z_{\langle \tilde{u}, \tilde{d} \rangle^A}^* \\ &> z_0 + |\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*| \\ &> z_0, \end{aligned}$$

which implies that  $\mathcal{E}_{A^*} \subset \mathcal{E}_1(Z^*)$ .

2.  $A$  contains the match  $\langle u_1, d_1 \rangle$ . Then  $A$  cannot contain the match  $\langle 1, 1 \rangle$ . Denote the set of such  $A$ 's as  $\mathcal{A}_2$ . Denote the set of  $E$ 's such that  $A^*$  has higher production than any  $A \in \mathcal{A}_2$  as

$$\mathcal{E}_2(Z^*) = \{E \mid S(A, E, Z^*) \leq S(A^*, E, Z^*), \forall A \in \mathcal{A}_2\},$$

which is equivalent to

$$\{E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq z_2^*(A), \forall A \in \mathcal{A}_2\},$$

where

$$z_2^*(A) = \left( \sum_{\langle u, d \rangle \in A^* \setminus \{A \cup \langle 1, 1 \rangle\}} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) + z_{\langle 1, 1 \rangle}^* - \gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*.$$

As above,  $\left( \sum_{\langle u, d \rangle \in A^* \setminus \{A \cup \langle 1, 1 \rangle\}} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) > z_0 + |\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*|$ . Thus

$$\begin{aligned} z_2^*(A) &> z_0 + |\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*| + z_{\langle 1, 1 \rangle}^* - \gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^* \\ &> z_0 + z_{\langle 1, 1 \rangle}^* \\ &> z_0, \end{aligned}$$

where the second inequality holds because either  $|\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*| - \gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^* = 0$  or because  $-\gamma_{\langle u_1, d_1 \rangle} z_{\langle u_1, d_1 \rangle}^*$  is positive and the third inequality holds because  $z_{\langle 1, 1 \rangle}^*$  was chosen to be positive before. These inequalities imply that  $\mathcal{E}_{A^*} \subset \mathcal{E}_2(Z^*)$ .

3.  $A$  contains neither  $\langle 1, 1 \rangle$  or  $\langle u_1, d_1 \rangle$ . Denote the set of such  $A$ 's as  $\mathcal{A}_3$ . Denote the set of  $E$ 's such that  $A^*$  has higher production than any  $A \in \mathcal{A}_3$  as

$$\mathcal{E}_3(Z^*) = \{E \mid S(A, E, Z^*) \leq S(A^*, E, Z^*), \forall A \in \mathcal{A}_3\},$$

which is equivalent to

$$\left\{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq \left( \sum_{\langle u, d \rangle \in A^* \setminus \{A \cup \langle 1, 1 \rangle\}} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) + z_{\langle 1, 1 \rangle}^*, \forall A \in \mathcal{A}_3 \right\}.$$

This implies  $\mathcal{E}_{A^*} \subset \mathcal{E}_3(Z^*)$  as above.

Henceforth  $\mathcal{E}_{A^*} \subset \mathcal{E}_i(Z^*)$  for  $i = 1, 2, 3$ , and

$$\Pr(A^* \mid Z^*) = \Pr[\mathcal{E}_1(Z^*) \cap \mathcal{E}_2(Z^*) \cap \mathcal{E}_3(Z^*)] \geq \Pr(\mathcal{E}_{A^*}) > \delta.$$

If  $\Pr(A^* \mid Z^*) \neq \hat{\Pr}(A^* \mid Z^*)$ , then the parameter  $\gamma_{\langle u_1, d_1 \rangle}$  is identified and we can move onto the identification of the other parameters, as discussed below. Thus we only consider the case  $\Pr(A^* \mid Z^*) = \hat{\Pr}(A^* \mid Z^*) > \delta$ . By Assumption 1 and Lemma 5, there exists  $z'_0$  such that  $\forall z \leq z'_0$ ,

$$\Pr(\tilde{S}(A, E) - \tilde{S}(A^*, E) < z, \forall A) < \delta. \quad (11)$$

Now let  $\check{Z} = (\check{z}_{\langle u, d \rangle})_{u, d \in N}$  where  $\check{z}_{\langle u, d \rangle} = z_{\langle u, d \rangle}^*$ , except that  $\check{z}_{\langle u_1, d_1 \rangle}(K) = K z_{\langle u_1, d_1 \rangle}^*$  and  $\check{z}_{\langle 1, 1 \rangle}(K) = K z_{\langle 1, 1 \rangle}^*$ . By the choice in 10,  $-(\gamma_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}(K))$  is negative and becomes more negative as  $K > 1$  increases. Therefore, there exists a sufficiently large  $K$  such that for all  $A \in \mathcal{A}_2$ ,

$$-(\gamma_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}(K)) < z'_0 - \left( \sum_{\langle u, d \rangle \in A^* \setminus A} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right),$$

which implies

$$\left( \sum_{\langle u, d \rangle \in A^* \setminus A} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) - (\gamma_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}(K)) < z'_0$$

and therefore

$$\begin{aligned} & \mathcal{E}_2(\check{Z}) \\ &= \left\{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq \left( \sum_{\langle u, d \rangle \in A^* \setminus A} \gamma_{\langle u, d \rangle} \check{z}_{\langle u, d \rangle} \right) - (\gamma_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}), \forall A \in \mathcal{A}_2 \right\} \\ &= \left\{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq \left( \sum_{\langle u, d \rangle \in A^* \setminus A} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) - (\gamma_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}), \forall A \in \mathcal{A}_2 \right\} \\ &\quad \subset \{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq z'_0, \forall A \in \mathcal{A}_2 \}, \end{aligned}$$

and by (11),

$$\Pr(\mathcal{E}_2(\check{Z})) < \Pr(E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq z'_0, \forall A \in \mathcal{A}_2) < \delta.$$

Therefore,

$$\Pr(A^* \mid \check{Z}) = \Pr(\mathcal{E}_1(\check{Z}) \cap \mathcal{E}_2(\check{Z}) \cap \mathcal{E}_3(\check{Z})) \leq \Pr(\mathcal{E}_2(\check{Z})) < \delta.$$

On the other hand, because the term  $-(\hat{\gamma}_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}(K))$  is positive, the set

$$\hat{\mathcal{E}}_2(\check{Z}) = \left\{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq \left( \sum_{\langle u, d \rangle \in A^* \setminus A} \hat{\gamma}_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) - (\hat{\gamma}_{\langle u_1, d_1 \rangle} \check{z}_{\langle u_1, d_1 \rangle}(K) - \check{z}_{\langle 1, 1 \rangle}(K)), \forall A \in \mathcal{A}_2 \right\}$$

becomes weakly larger as  $K$  increases. Therefore for  $K > 1$ ,  $\hat{\mathcal{E}}_2(Z^*) \subset \hat{\mathcal{E}}_2(\check{Z})$ . We also have

$\hat{\mathcal{E}}_3(Z^*) \subset \hat{\mathcal{E}}_3(\check{Z})$ , because

$$\mathcal{E}_3(\check{Z}) = \left\{ E \mid \tilde{S}(A, E) - \tilde{S}(A^*, E) \leq \left( \sum_{\langle u, d \rangle \in A^* \setminus \{A \cup \langle 1, 1 \rangle\}} \gamma_{\langle u, d \rangle} z_{\langle u, d \rangle}^* \right) + \check{z}_{\langle 1, 1 \rangle}(K), \forall A \in \mathcal{A}_3 \right\},$$

whose probability increases with  $K$ ;  $\hat{\mathcal{E}}_1(\check{Z}) = \hat{\mathcal{E}}_1(Z^*)$ , because  $z_1(A)$  does not contain the production of the matches of  $\langle 1, 1 \rangle$  and  $\langle u_1, d_1 \rangle$ . Therefore,

$$\begin{aligned} \hat{\Pr}(A^* \mid \check{Z}) &= \hat{\Pr}(\hat{\mathcal{E}}_1(\check{Z}) \cap \hat{\mathcal{E}}_2(\check{Z}) \cap \hat{\mathcal{E}}_3(\check{Z})) \\ &\geq \hat{\Pr}(\hat{\mathcal{E}}_1(Z^*) \cap \hat{\mathcal{E}}_2(Z^*) \cap \hat{\mathcal{E}}_3(Z^*)) \\ &\geq \delta. \end{aligned}$$

Therefore,  $\hat{\Pr}(A^* \mid \check{Z}) \geq \delta$  and  $\Pr(A^* \mid \check{Z}) < \delta$ .

Therefore,  $\gamma_{\langle u_1, d_1 \rangle} = \hat{\gamma}_{\langle u_1, d_1 \rangle}$  if  $u_1 = 1$  or  $d_1 = 1$ . Now we can show that  $\gamma_{\langle u, d \rangle} = \hat{\gamma}_{\langle u, d \rangle}$  for any  $\langle u, d \rangle$  by slightly modifying the above proof steps. Suppose  $\gamma_{\langle u, d \rangle} \neq \hat{\gamma}_{\langle u, d \rangle}$ , and without loss of generality,  $u \neq 1$ . Then by the previous argument, we have  $\gamma_{\langle 1, d \rangle} = \hat{\gamma}_{\langle 1, d \rangle}$ . Then in the above proof, we can replace  $z_{\langle 1, 1 \rangle}$  with  $\gamma_{\langle 1, d \rangle} z_{\langle 1, d \rangle}$ , have  $A^*$  contain the match  $\langle 1, d \rangle$  and generate a similar contradiction. For the case of  $\gamma_{\langle 1, 1 \rangle} = -1$ , simply reverse the sign of  $z_{\langle 1, 1 \rangle}$  and the proof is identical.  $\square$

Once the parameter matrix  $\Gamma$  is identified, we can then constructively trace  $F(C)$  by a variant of Theorem 2.

## A.5 Proof of Theorem 4

Condition on  $X$ . We let  $w(E)$  be the matrix of unobserved production levels

$$w(E) = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & f_{\langle 2, 2 \rangle} & \dots & f_{\langle 2, N \rangle} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f_{\langle N, 2 \rangle} & \dots & f_{\langle N, N \rangle} \end{pmatrix},$$

where we define

$$f_{\langle u, d \rangle} = e_u \cdot e_d.$$

Let  $\tilde{S}(A, E) = \sum_{u=1}^N f_{\langle u, A(u) \rangle}$  and  $S(A, E, Z) = \tilde{S}(A, E) + \sum_{u=1}^N z_u \cdot z_d$ . Let  $A_1 = \{\langle 1, 1 \rangle, \dots, \langle N, N \rangle\}$ . Also define  $\bar{S}(A, E) = \tilde{S}(A, E) - \tilde{S}(A_1, E)$  and the vector  $\bar{y}(E) = (\bar{S}(A, E))_{A \neq A_1}$ . First we show that there is a one-to-one and onto relationship between  $\bar{y}(E)$  and  $E$ . This amounts to showing

$E_1 = E_2$  if and only if  $\bar{y}(E_1) = \bar{y}(E_2)$ . The “only if” direction is trivial. The following two lemmas are dedicated to showing the “if” part of the assertion.

**Lemma 6.** *If  $\tilde{S}(A, E_1) = \tilde{S}(A, E_2)$  for all  $A$ , then  $E_1 = E_2$ .*

*Proof.* Let  $f^1$  index the unobserved production for  $E_1$  and let  $f^2$  index the unobserved production for  $E_2$ . Recall that  $f_{\langle u, d \rangle} = e_u \cdot e_d$ . We start on the diagonal elements. For assignment  $(\langle 1, 1 \rangle, \langle 2, 2 \rangle, \dots, \langle N, N \rangle)$ ,

$$\sum_{u=2}^N f_{\langle u, u \rangle}^1 = \sum_{u=2}^N f_{\langle u, u \rangle}^2, \quad (12)$$

as the missing first term in each sum is 0 as  $e_1$  is normalized to be 0. For assignment  $(\langle 1, k \rangle, \langle 2, 2 \rangle \dots, \langle k, 1 \rangle, \dots, \langle N, N \rangle)$ , we have

$$\sum_{\substack{u=2 \\ u \neq k}}^N f_{\langle u, u \rangle}^1 = \sum_{\substack{u=2 \\ u \neq k}}^N f_{\langle u, u \rangle}^2, \quad (13)$$

again as the matches  $\langle 1, k \rangle$  and  $\langle k, 1 \rangle$  have zero non- $z$  production by the location normalization that  $e_1 = 0$  for upstream and downstream firms. Subtract equation (13) from equation (12), which gives us  $f_{\langle k, k \rangle}^1 = f_{\langle k, k \rangle}^2$ . The previous argument can be repeated for all  $k$ , giving  $f_{\langle k, k \rangle}^1 = f_{\langle k, k \rangle}^2 \forall k = 1, \dots, N$ .

Next we induct on  $N$ . When  $N = 2$ , the only nonzero term in  $w(E)$  is  $f_{\langle 2, 2 \rangle}$ , which we just showed was identical between  $E_1$  and  $E_2$ . Now assume for any  $N - 1$ -by- $N - 1$  submarket,  $w(E_1) = w(E_2)$  only if  $E_1 = E_2$ . In an  $N$ -by- $N$  market, we know the element  $f_{\langle N, N \rangle}$  is the same. For arbitrary elements  $f_{\langle N, d \rangle}$  of the last row and  $f_{\langle u, N \rangle}$  of the last column, we want to show that those elements are identical between  $E_1$  and  $E_2$ . If we can show this for an element of the last row, a symmetric argument will hold for an element of the last column. Focusing on  $f_{\langle N, d \rangle}$ , consider the assignment

$$(\langle N - 1, 1 \rangle, \langle N - 2, 2 \rangle, \dots, \langle N - d + 1, d - 1 \rangle, \langle N, d \rangle, \langle N - d, d + 1 \rangle, \dots, \langle 1, N \rangle),$$

which has the sum of unobserved production

$$0 + f_{\langle N-2, 2 \rangle} + \dots + f_{\langle N-d+1, d-1 \rangle} + f_{\langle N, d \rangle} + f_{\langle N-d, d+1 \rangle} + \dots + 0.$$

All terms in the sum except for  $f_{\langle N, d \rangle}$  are equal across  $E_1$  and  $E_2$  by the induction hypothesis and the location normalization. Because the sums of non- $z$  production are equal across assignments,  $f_{\langle N, d \rangle}^1 = f_{\langle N, d \rangle}^2, \forall d$ . Thus the induction is complete and  $w(E_1) = w(E_2)$ .

With the scale normalization that  $e_u = 1$  for  $u = 2$  and  $f_{\langle u, d \rangle} = e_u \cdot e_d$ , we have  $e_2 \cdot e_d^1 = e_2 \cdot e_d^2$ , which implies that  $e_d^1 = e_d^2$ . By comparing every term that involves  $f_{\langle u, d \rangle}, u \neq 1, 2$ , we have  $e_u^1 = e_u^2$ . Therefore  $E_1 = E_2$ .  $\square$

Next we prove the following lemma.

**Lemma 7.** *Let  $A_1$  be the diagonal assignment as defined at the beginning of the section.*

$$\tilde{S}(A_2, E_1) - \tilde{S}(A_1, E_1) = \tilde{S}(A_2, E_2) - \tilde{S}(A_1, E_2), \forall A_2 \neq A_1, \quad (14)$$

*if and only if  $E_1 = E_2$ .*

*Proof.* The “if” direction is trivial. For the “only if” direction, by re-arranging terms, we have

$$\tilde{S}(A_2, E_1) - \tilde{S}(A_2, E_2) = \tilde{S}(A_1, E_1) - \tilde{S}(A_1, E_2).$$

We can define  $D$  to be equal to the value of the right side, which is not a function of  $A_2$ . Given the implication of the location normalizations that  $f_{\langle 1,1 \rangle} = 0$ , we have

$$\tilde{S}(A_1, E_1) - \tilde{S}(A_1, E_2) = \sum_{u=2}^N f_{\langle u,u \rangle}^1 - \sum_{u=2}^N f_{\langle u,u \rangle}^2 = D, \quad (15)$$

where the superscripts index  $E_1$  and  $E_2$ . Consider the assignment  $A_2 = (\langle 1, k \rangle, \langle 2, 2 \rangle \cdots, \langle k, 1 \rangle, \cdots, \langle N, N \rangle)$ . Under the implication of the location normalizations that  $f_{\langle 1,k \rangle} = f_{\langle k,1 \rangle} = 0$ , we have

$$\tilde{S}(A_2, E_1) - \tilde{S}(A_2, E_2) = \sum_{\substack{u=2 \\ u \neq k}}^N f_{\langle u,u \rangle}^1 - \sum_{\substack{u=2 \\ u \neq k}}^N f_{\langle u,u \rangle}^2 = D \quad (16)$$

Comparing (15) and (16), we have that  $f_{\langle k,k \rangle}^1 = f_{\langle k,k \rangle}^2$ . Varying the argument for each  $k$  shows that  $f_{\langle k,k \rangle}^1 = f_{\langle k,k \rangle}^2$  for all  $k$  as functions of  $X$ . Therefore  $\sum_{u=2}^N f_{\langle u,u \rangle}^1 = \sum_{u=2}^N f_{\langle u,u \rangle}^2$  and  $D = 0$ . Therefore  $\tilde{S}(A_2, E_1) = \tilde{S}(A_2, E_2) \forall A_2$ . By Lemma 6, this implies that  $E_1 = E_2$ .  $\square$

Now we are ready to prove Theorem 4.

*Proof.* Continue to condition on  $X$ . By Lemmas 7, there is a one-to-one and onto relationship between  $\bar{y}(E)$  and  $E$ . Thus it suffices to show that the distribution  $H(\bar{y}(E))$  is identified. The rest of the proof is parallel to the proof of Lemma 4 and Theorem 2.

Suppose  $E^*$  gives  $\bar{y}(E^*)$ . Choose  $Z^*$  such that

$$z_u^* = -e_u^*, \quad z_d^* = e_d^*.$$



Then  $S(A, E^*, Z^*) = 0 \forall A$ . At  $\bar{y}(E^*)$ , the CDF has the value of

$$\begin{aligned}
H(\bar{y}(E^*)) &= \Pr(\tilde{S}(A_2, E) - \tilde{S}(A_1, E) \leq \tilde{S}(A_2, E^*) - \tilde{S}(A_1, E^*), \dots) \\
&= \Pr(S(A_2, E, Z^*) - S(A_1, E, Z^*) \leq S(A_2, E^*, Z^*) - S(A_1, E^*, Z^*), \dots \mid Z^*) \\
&= \Pr(S(A_2, E, Z^*) - S(A_1, E, Z^*) \leq 0, \dots \mid Z^*) \\
&= \Pr(A_1 \mid Z^*).
\end{aligned}$$

The first equality is the definition of a CDF; the second equality adds the production from observables in  $Z^*$  for each assignment  $A$  and the base assignment  $A_1$  to both sides of the inequalities; the third equality uses the particular choice  $Z^*$  and its implication  $S(A, E^*, Z^*) = 0 \forall A$ ; finally, the fourth equality uses the fact that assignment  $A_1$  occurs when the differences between the productions of each other assignment  $A$  and of  $A_1$  is nonpositive. Therefore the cumulative distribution function  $H(\bar{y}(E^*))$  can be traced by  $\Pr(A_1 \mid Z^*)$ . By Lemma 7, applying a bijective change of variables gives us the distribution of  $E^*$ . Therefore the distribution  $G(E)$  is identified.  $\square$

## References

- Ackerberg, Daniel A. and Maristella Botticini**, “Endogenous Matching and the Empirical Determinants of Contract Form,” *Journal of Political Economy*, 2002, 110 (2002), 564–591.
- Baccara, Mariagiovanna, Ayse Imrohoroglu, Alistair Wilson, and Leeat Yariv**, “A Field Study on Matching with Network Externalities,” *American Economic Review*, forthcoming.
- Bajari, Patrick, Han Hong, and Stephen Ryan**, “Identification and Estimation of a Discrete Game of Complete Information,” *Econometrica*, September 2010, 78 (5), 1529–1568.
- Becker, Gary S.**, “A Theory of Marriage: Part I,” *Journal of Political Economy*, July–August 1973, 81 (4), 813–846.
- Beresteanu, Arie, Ilya Molchanov, and Francesca Molinari**, “Sharp Identification Regions in Models with Convex Moment Predictions,” *Econometrica*, 2011, 79 (6), 1785–1821.
- Berry, S. and E. Tamer**, “Identification in models of oligopoly entry,” in Whitney Newey Richard K. Blundell and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress*, Cambridge University Press, 2006, chapter 2, p. 46.
- 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.
- Bresnahan, Timothy F. and Peter C. Reiss**, “Entry and Competition in Concentrated Markets,” *The Journal of Political Economy*, October 1991, 99 (5), 977–1009.
- Brock, W.A. and S.N. Durlauf**, “Identification of binary choice models with social interactions,” *Journal of Econometrics*, 2007, 140 (1), 52–75.
- Brown, James N. and Harvey S. Rosen**, “On the Estimation of Structural Hedonic Price Models,” *Econometrica*, May 1982, 59 (3), 765–768.
- Chen, Jiawei**, “Two-Sided Matching and Spread Determinants in the Loan Market,” May 2009. working paper.
- Chiappori, P.A., R.J. McCann, and L.P. Nesheim**, “Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness,” *Economic Theory*, 2010, 42 (2), 317–354.
- Chiappori, Pierre-André, Alfred Galichon, and Bernard Salanié**, “The Roommate Problem Is More Stable Than You Think,” 2012. Columbia University working paper.
- , **Bernard Salanié, and Yoram Weiss**, “Partner Choice and the Marital College Premium,” 2010. Columbia University working paper.
- Choo, Eugene and Aloysius Siow**, “Who Marries Whom and Why,” *The Journal of Political Economy*, November 2006, 114 (1), 175–201.
- Christakis, Nicholas, James Fowler, Guido W. Imbens, and Karthik Kalyanaraman**, “An Empirical Model for Strategic Network Formation,” 2010. working paper.
- Ciliberto, Federico and Elie Tamer**, “Market Structure and Multiple Equilibria in Airline Markets,” *Econometrica*, November 2009, 77 (6), 1791–1828.
- Crawford, Gregory S. and Ali Yurokoglu**, “The Welfare Effects of Bundling in Multichannel Television Markets,” *American Economic Review*, forthcoming.
- Dagsvik, John K.**, “Aggregation in Matching Markets,” *International Economic Review*, February 2000, 41 (1), 27–57.

- de Paula, Aureo and Xun Tang**, “Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information,” *Econometrica*, 2012, 80 (1), 143–172.
- Echenique, Federico, SangMok Lee, Matthew Shum, and M. Bumin Yenmez**, “The Revealed Preference Theory of Stable and Extremal Stable Matchings,” 2011. Caltech working paper.
- Eeckhout, Jan and Philipp Kircher**, “Identifying Sorting – In Theory,” *Review of Economic Studies*, 2011, 78, 872–906. University Pompeu Fabra working paper.
- Ekeland, Ivar, James J. Heckman, and Lars Nesheim**, “Identification and Estimation of Hedonic Models,” *Journal of Political Economy*, 2004, 112 (S1), 60–109.
- 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 Multinomial Choice Models,” 2012. University of Michigan working paper.
- **and Kyoo il Kim**, “A Simple Nonparametric Approach to Estimating the Distribution of Random Coefficients in Structural Models,” 2011. University of Michigan working paper.
- Galichon, A. and M. Henry**, “Set identification in models with multiple equilibria,” *The Review of Economic Studies*, 2011, 78 (4), 1264–1298.
- Galichon, Alfred and Bernard Salanie**, “Matching with Trade-offs: Revealed Preferences over Competing Characteristics,” 2010. Ecole Polytechnique working paper.
- Gautier, E. and Y. Kitamura**, “Nonparametric Estimation in Random Coefficients Binary Choice Models,” 2011. CREST working paper.
- Gordon, Nora and Brian Knight**, “A spatial merger estimator with an application to school district consolidation,” *Journal of Public Economics*, 2009, 93, 752–765.
- Graham, Bryan S.**, “Econometric methods for the analysis of assignment problems,” in A. Bisin J. Benhabib and M. Jackson, eds., *Handbook of Social Economics*, 2010.
- Hatfield, John William and Paul R. Milgrom**, “Matching with Contracts,” *The American Economic Review*, September 2005, 95 (4), 913–935.
- , **Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp**, “Stability and Competitive Equilibrium in Trading Networks,” January 2011. Stanford University working paper.

- Heckman, James J., Rosa L. Matzkin, and Lars Nesheim**, “Nonparametric Identification and Estimation of Nonadditive Hedonic Models,” *Econometrica*, September 2010, 78 (5), 1569–1591.
- Hitsch, Günter, Ali Hortaçsu, and Dan Ariely**, “Matching and Sorting in Online Dating,” *American Economic Review*, 2010.
- Ho, Katherine**, “Insurer-provider networks in the medical care market,” *American Economic Review*, 2009, 99 (1), 393–430.
- 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. Washington University in St Louis working paper.
- Lewbel, Arthur**, “Semiparametric latent variable model estimation with endogenous or mismeasured regressors,” *Econometrica*, 1998, 66 (1), 105–121.
- , “Semiparametric Qualitative Response Model Estimation with Unknown Heteroscedasticity or Instrumental Variables,” *Journal of Econometrics*, 2000, 97 (1), 145–177.
- Logan, J.A., P.D. Hoff, and M.A. Newton**, “Two-sided estimation of mate preferences for similarities in age, education, and religion,” *Journal of the American Statistical Association*, 2008, 103 (482), 559–569.
- Manski, Charles F.**, “Identification of Binary Response Models,” *Journal of the American Statistical Association*, September 1988, 83 (403), 729–738.
- Matzkin, Rosa L.**, “Heterogeneous Choice,” in W. Newey and T. Persson, eds., *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society*, Cambridge University Press., 2007.
- Mazzeo, Michael J.**, “Product Choice and Oligopoly Market Structure,” *The RAND Journal of Economics*, 2002, 33 (2), 221–242.
- Menzel, Konrad**, “Robust Decisions for Incomplete Structural Models of Social Interactions,” 2011. New York University working paper.
- Park, Minjung**, “An Empirical Two-Sided Matching Model of Acquisitions: Understanding Merger Incentives and Outcomes in the Mutual Fund Industry,” September 2008. University of California, Berkeley 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.
- Seim, Katja**, “An empirical model of firm entry with endogenous product-type choices,” *The RAND Journal of Economics*, 2006, *37* (3), 619–640.
- Shapley, Lloyd S. and Martin Shubik**, “The assignment game I: the core,” *International Journal of Game Theory*, 1972, *1*, 111–130.
- Siow, Aloysius**, “Testing Becker’s Theory of Positive Assortative Matching,” Working Papers tecipa-356, University of Toronto, Department of Economics April 2009.
- Sørensen, Morten**, “How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital,” *Journal of Finance*, December 2007, *LXII* (6), 2725–2762.
- Sotomayor, Marilda**, “The lattice structure of the set of the stable outcomes of the multiple partners assignment game,” *International Journal of Game Theory*, November 1999, *28* (4), 567–583.
- Tamer, Elie**, “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria,” *Review of Economic Studies*, 2003, *70* (1), 147–165.
- Thompson, T. Scott**, “Identification of Semiparametric Discrete Choice Models,” 1989. University of Minnesota Discussion Paper 249.
- Weese, Eric**, “Political Mergers as Coalition Formation: Estimates from Japanese Municipal Mergers,” 2010. Yale University working paper.
- 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.