Estimating Matching Games with Profit and Price Data

Jeremy T. Fox
Rice University & NBER

Amir Kazempour
Rice University

Xun Tang
Rice University

February 2024

Abstract

Empirical methods for transferable-utility matching games have previously been developed using the key outcome of the matches formed in equilibrium. We explore identification and estimation of match production functions and agent valuation functions using data on two additional outcomes of such matching games: monetary transfers (prices) and profits. We provide identification results for nonparametric models for the case of data on profits and for more parametric models for the case of data on prices. We provide estimators paralleling the identification results for both profit data and price data. Importantly, our identification results allow for agents to have valuations defined over the unmeasured characteristics of potential partners.

1 Introduction

This paper studies the identification and estimation of aspects of the payoffs of agents in two-sided matching games using data on prices or profits in addition to data on matches. We first describe some background on the empirical use of matching games without price and profit data and then move onto our results incorporating price or profit data.

Matching games make predictions about which of a heterogeneous group of agents on one side of a matching market will match with particular agents on the other side of the matching market. Transferable-utility matching games are a particular type of matching game where matched agents exchange monetary transfers and those transfers enter additively separably into the payoffs of agents for matches. Since at least Becker (1973), economists have used transferable-utility matching games to make empirical predictions about the matching patterns involving the heterogeneous agents on both sides of a matching market. A famous theoretical result is sometimes used. If the years of schooling of men and women are the only agent characteristics entering the production of a match, high-schooling men marry high-schooling women and low-schooling men marry low-schooling women if the schooling levels of men and women are complements in the production of a match.

More recently, an influential literature developed methods to structurally estimate the parameters of the match production function in transferable utility matching games (e.g., Choo and Siow 2006; Fox 2010; Dupuy and Galichon 2014; Chiappori, Salanié, and Weiss 2017; Fox, Yang, and Hsu 2018; Fox 2018; Galichon and Salanié 2022). The methods use data on who matches with whom as the dependent variable, the outcome of the matching game observed by the researcher. Moving past the analysis of scalar agent characteristics in Becker (1973), the independent variables in the data are vectors of measured agent characteristics as well as, sometimes, match-specific characteristics. There are some limitations to the identification results in this literature. One of the limitations is that usually some notion of complementarities is the only aspect of match production that is identified without using data on agents who are unmatched. Data on unmatched agents are not always available. A further limitation is that at best match production functions can be identified only up to location and scale normalizations.

The outcome to a transferable-utility matching game has additional components. An outcome to a transferable-utility matching game is comprised of the matches (who matches with whom) and, importantly, the monetary transfers, which we often call prices, exchanged between the matched parties. Together, the match and price outcomes can be used to compute the equilibrium profit (or utility) of each agent in a matching market if the valuation functions are known. There exist many datasets where the researcher observes data on who matches with whom in addition to data on the prices exchanged between matched agents or the profits agents that have. Both price and profit data are measured in monetary units, such as dollars.

The current paper studies the identification and estimation of aspects of the match production function as well as the valuation functions of the two sides of the matching market using data on matches as well as one or both of the continuous, monetary outcomes of prices and profits. The addition of a mostly new type of dependent-variable data has the potential to resolve some of the limitations of identification using data on matches only discussed above. It is natural to expect that using monetary outcome data will allow the identification of the scale of payoff-related functions (match production and match valuations) in monetary units. Further, our results explore to what degree price or profit data can identify the level of match production and match valuations, rather than just some notion of complementarities, without relying on data on unmatched agents.

We work with outcome and agent characteristic data sampled from one or a finite number of one-to-one, two-sided, continuum matching markets. By a continuum, we mean that each matching market in truth has an infinite number of agents. Continuum matching games are also studied in the majority of the structural empirical literature on estimating transferable-utility matching games using outcome data on matches only (e.g., Choo and Siow 2006; Fox 2010; Dupuy and Galichon 2014; Chiappori, Salanié, and Weiss 2017; Fox 2018; Galichon and Salanié 2022). While our paper and this prior empirical literature on continuum matching games both study one-to-one, two-sided transferable-utility matching games, there are some key differences in the continuum models in the matches-only literature and the models we explore. First, we focus on the case where agents on both of the two sides of the market each have a vector of measured, continuous characteristics. The literature on using outcome data on matches only often focuses on on discrete characteristics, with Fox (2010), Dupuy and Galichon (2014) and to some degree Fox (2018) being notable exceptions. Second, our treatment of variables that are unmeasured in the data (the econometric error terms) is quite different. Most of our identification results assume each agent has a scalar, unmeasured agent characteristic. This unmeasured agent characteristic is valued by agents on the other side of the market and can shift the valuations of the agent whose characteristic it is. By contrast, the prior literature on estimating continuum matching games using matches only relies on an model where unmeasured variables reflect preferences over the measured characteristics of partners. For example in Choo and Siow (2006), who study marriage, men and women have the measured characteristic of schooling and men have logit shocks over female schooling levels and women have logit shocks over male schooling levels. In the prior literature on continuum matching games, there are no unmeasured agent characteristics that are valued by the other side of the market. In the sense that we allow such unmeasured agent characteristics, our continuum matching game shares features with the finite-agent matching games analyzed in Fox, Yang, and Hsu (2018).

As our paper studies matching games using match, price and profit data, we need to consider several different data schemes in our paper. In labor markets, it is common in linked employee-employer data to observe the salary of a worker in addition to characteristics of the worker and his or her employer. This salary is a monetary transfer and hence is the price data we would use in our analysis. If the worker is assumed to have no non-wage valuations over firms, a common assumption in models of the labor market, then this salary could also be modeled as the profit of the worker. In business-to-business markets, it is common to observe the price of a transaction between a buyer and a seller. It is less common but still possible to observe the profit of either or both of the buyer and seller, say from accounting data. We generally assume that match data are available, although one of our results using profit data does not use information on the characteristics of matched partners.

While we consider many combinations of data availability and restrictions on match production and valuation functions, there are some broad differences between the types of results we present for profit data and for price data. For profit data, our identification and estimation results treat the unknown match production function nonparametrically, meaning it is not specified up to a finite vector of unknown parameters. For estimation, these nonparametric identification results inspire computationally simple estimators. For price data, our recommended estimation approach is parametric and reasonably computationally intense, as it involves computing the equilibrium to the matching game inside the evaluation of a likelihood function by an optimization routine. The intuition for the more straightforward identification results for profit data is the profits are the equilibrium utilities of agents and correspond closely with agents' objectives when making matches in the game. By contrast, prices are equilibrium transfers between two matched agents and shift the payoffs of the agents for particular partners, rather than being the realized payoffs themselves, as in the case of data on profits.

Each section below describes the results on matching games and results from prior papers in econometrics that we rely on for each of our identification results and our estimation procedures. For the case of profits data, we build on methods that have studied non-separable econometric models (e.g., Matzkin 2003). For the case of prices data, we use methods for simulated likelihoods where the outcome data is continuous, as in our data schemes (e.g., Fermanian and Salanié 2004).

In addition to our results on identification and estimation, we present one result on matching theory that may be of independent interest. Theorem 1 shows sufficient conditions for the random variables corresponding to the measured characteristics of agents matched to each other in equilibrium to be statistically independent of the random variables corresponding to the unmeasured characteristics of agents matched to each other in equilibrium. As the conditions in Theorem 1 may at first seem to be appealing starting places for identification analysis, the new theorem shows that weakening the conditions using a fully non-separable model may be necessary to fit typical datasets, which contain many matches between agents with measured characteristics such that the model needs econometric error terms, in our model the unmeasured agent characteristics, to fit the match patterns in the data.

The literature that is closest to our study of price (not profit) data from transferable-utility matching games goes by a different name: hedonics. In a hedonics model, buyers and sellers transact both for monetary transfers and for other, endogenously determined, non-price product characteristics. Think of the motivating example of housing. Buyers are families who match with sellers, who are construction builders. The notion of a match in the housing market would include the price of the house, the transfer in the model, and some non-price characteristics of the house, like how big the house is. If this optimal choice of non-price characteristics is concentrated out, under appropriate restrictions a hedonics model is a transferable utility matching game between buyers and sellers. Previous papers on identification in hedonics models use data on prices in addition to non-price characteristics of the product (e.g., Ekeland, Heckman, and Nesheim 2004; Bajari and Benkard 2005; Heckman, Matzkin, and Nesheim 2010). A key restriction used in previous analyses of identification in hedonics models has been that agents have valuation functions defined over the non-price characteristics and not the measured or unmeasured characteristics of their match partners. In housing, families have valuations defined over the measured characteristics of the house and not the measured or unmeasured characteristics of the contractor that builds the house. The unmeasured variables in previous identification results in the hedonics literature are preference shifters over the measured, non-price characteristics of the product in question, like a newly built house. Therefore, the treatment of unmeasured variables as shifters of preferences over measured characteristics is to some degree reminiscent of the treatment of unmeasured variables in the literature on the identification in continuum matching games using outcome data on matches only, as described earlier (e.g., Choo and Siow 2006). Our model allows agents to have valuations defined over the unmeasured characteristics of partners.

Section 2 describe the one-to-one, two-sided, continuum matching game that we study. Section 3 considers identification under profit data, using nonparametric models. Section 4 considers identification under price data, focusing with provided motivation on parametric models. Section 5 then discusses estimation with price data and with profit data in separate subsections. We plan to add an empirical application.

2 The Model

We first introduce some basic notations and definitions before discussing the existence and uniqueness of the equilibrium.

Notations and Definitions

Agents belong to either the *upstream* side or the *downstream* side of the market. An upstream firm is characterized by a vector of types denoted by $\tilde{\boldsymbol{X}} \equiv (\boldsymbol{X}, \boldsymbol{\varepsilon}) \in \tilde{\mathcal{X}}$ where $\boldsymbol{x} \in \mathcal{X}$ is a d_x-dimension vector of the *observable* or *measured* characteristics and $\boldsymbol{\varepsilon} \in \mathcal{E}$ is a d_ϵ-dimension vector of the *unobservable* or *unmeasured* characteristics. Similarly, a downstream firm is characterized by $\tilde{y} \equiv (\boldsymbol{y}, \boldsymbol{\eta}) \in \tilde{\mathcal{Y}}$, where $\boldsymbol{y}$ is a d_y-dimension vector of observable characteristics and $\boldsymbol{\eta} \in \mathcal{H}$ is a d_η-dimension vector of unobservable characteristics. In the following sections, unless clearly stated otherwise, we assume that unobserved characteristics of the upstream and downstream firms are both scalars (i.e., $d_x = d_y = 1$), and denote them with ε and η. We will explicitly specify where the results can be extended to allow for vector unobservable characteristics. The upstream firm's types are distributed as $F^u(\cdot, \cdot)$, whereas the downstream firm's types are distributed as $F^d(\cdot, \cdot)$. Throughout the paper, we assume that the type distributions $F^u(\cdot, \cdot)$and $F^d(\cdot, \cdot)$ are continuous and atom-less distributions over $\mathbb{R}^{d_x + d_\varepsilon}$ and $\mathbb{R}^{d_y + d_\eta}$. The mass of firms on each side of the market is normalized to 1.

A *match* between an upstream firm $\tilde{\boldsymbol{x}}$ and a downstream firm $\tilde{\boldsymbol{y}}$ generates (pairwise) match production

$$\Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}\right) = \Phi\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}\right),$$

where the match production function $\Phi : \mathbb{R}^{d_x+d_\epsilon+d_y+d_\eta} \mapsto \mathbb{R}$ is a model primitive. It is worth noting that we do not specify the firm's payoff in a match because it is an equilibrium outcome of the model.

A *feasible matching,* also known as a *coupling* in the optimal transport literature, determines which firm on each side of the market is matched to another firm on the other side of the market. In a finite setting, a matching can be characterized by the finite set of all matches. In a continuous setting, where there is an uncountable number of matches, a matching is defined as a probability measure μ over $\mathcal{X} \times \mathcal{Y}$ determining the frequency of a match between any pair of upstream and downstream firm types $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}})$. However, not every probability measure on $\mathcal{X} \times \mathcal{Y}$ is a feasible matching. Specifically, the matching μ constitutes a feasible matching if and only if the marginals ofμ coincide with the distributions $F^u(\cdot,\cdot)$ and $F^d(\cdot,\cdot)$. This is analogous to the condition that each firm is matched only once in a finite one-to-one matching model, or matched less than or equal to a quota in a finite many-to-many matching model. The set of all feasible matchings between $F^u(\cdot,\cdot)$ and $F^d(\cdot,\cdot)$ is denoted by $\mathcal{M}\left(F^u, F^d\right)$.

A feasible matching where the downstream partner of $(\boldsymbol{x}, \boldsymbol{\varepsilon})$ is a deterministic function of $(\boldsymbol{x}, \boldsymbol{\varepsilon})$ and similarly the upstream partner of $(\boldsymbol{y}, \boldsymbol{\eta})$ is a deterministic function of $(\boldsymbol{y}, \boldsymbol{\eta})$ is called a *pure matching.* In a pure matching, let $T^u : \mathbb{R}^{d_x} \times \mathbb{R}^{d_\epsilon} \mapsto \mathbb{R}^{d_y}$ denote the mapping from the space of upstream types $\tilde{\mathcal{X}}$ to the space $\mathcal{Y}$ of the observed downstream types. That is, for $(\boldsymbol{x}, \boldsymbol{\varepsilon}) \in \mathbb{R}^{d_x+d_\epsilon}$, the observed partner type $\boldsymbol{y}$ is given by $T^u(\boldsymbol{x}, \boldsymbol{\varepsilon})$, where $\boldsymbol{y} \in \mathbb{R}^{d_y}$ is the unique value such that $\mu(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}) > 0$ for some $\boldsymbol{\eta} \in \mathbb{R}^{d_\eta}$. The mapping $T^u(\boldsymbol{x}, \boldsymbol{\varepsilon})$ is the observed type of the unique partner of $(\boldsymbol{x}, \boldsymbol{\varepsilon})$-firm. Similarly, we define the mapping $S^u : \mathbb{R}^{d_x} \times \mathbb{R}^{d_\epsilon} \mapsto \mathbb{R}^{d_\eta}$ given by $S^u(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \boldsymbol{\eta}$ where $\boldsymbol{\eta} \in \mathbb{R}^{d_\eta}$ is the unique value such that $\mu(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}) > 0$ for some $\boldsymbol{y} \in \mathbb{R}^{d_y}$. The mappings $T^d(\boldsymbol{y}, \boldsymbol{\eta})$ and $S^d(\boldsymbol{y}, \boldsymbol{\eta})$ are s the observed and unobserved type of the unique matching partner of the downstream firm $(\boldsymbol{y}, \boldsymbol{\eta})$.

A *pairwise stable pure equilibrium* in the continuous two-sided matching model is a pure matching that is characterized by $(T^u(\cdot), S^u(\cdot))$ and a pair of equilibrium profit functions $\left(\pi^u, \pi^d\right)$, where $\pi^u : \mathbb{R}^{d_x+d_\varepsilon} \mapsto \mathbb{R}$ and $\pi^d : \mathbb{R}^{d_y+d_\eta} \mapsto \mathbb{R}$ are such that

$$\pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) + \pi^d\left(\boldsymbol{y}, \boldsymbol{\eta}\right) \geq \Phi\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}\right), \quad \text{if } \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{\eta} \end{pmatrix} \neq \begin{pmatrix} T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) \\ S^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) \end{pmatrix}, \tag{1}$$

and

$$\pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) + \pi^d\left(\boldsymbol{y}, \boldsymbol{\eta}\right) = \Phi\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}\right), \quad \text{if } \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) \\ S^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) \end{pmatrix}. \tag{2}$$

Equation (1) says that the sum of the equilibrium profits for any pair of upstream and downstream firms should be at least as much as the potential production that they would generate if they were in a match. If (1) is violated for some $(\boldsymbol{x}, \boldsymbol{\varepsilon}) \in \tilde{\mathcal{X}}$ and $(\boldsymbol{y}, \boldsymbol{\eta}) \in \tilde{\mathcal{Y}}$ such that $\boldsymbol{y} \neq T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)$ and/or $\boldsymbol{\eta} \neq S^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)$, then they can deviate from their equilibrium matches under (T^u, S^u) and form a match with each other by dividing the production net of their equilibrium profits, i.e., $\Phi\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}\right) - \pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) + \pi^d\left(\boldsymbol{y}, \boldsymbol{\eta}\right) > 0$, equally so that they are both strictly better off compared to their equilibrium payoffs. Equation (2) says that upstream and downstream firms in a match share the match production between themselves according to the equilibrium profit functions $\pi^u\left(\cdot, \cdot\right)$ and $\pi^d\left(\cdot, \cdot\right)$.

Existence and Uniqueness

The existence and uniqueness of a pure equilibrium matching and equilibrium profits follow from a variation of the Monge-Kantorovich theorem in Carlier (2003). We first state the assumptions on the model primitives to ensure the existence and uniqueness of the pure equilibrium matching—see Galichon (2016, p. 76).

Assumption 1. *(i) (Differentiability) The match production function $\Phi\left(\cdot, \cdot, \cdot, \cdot\right)$ is twice continuously differentiable in all its arguments and for every compact set $K_y \subseteq \tilde{\mathcal{Y}}$, there exists a constant $c_{K_y} > 0$ such that for every $\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2 \in \tilde{\mathcal{X}}$,*

$$\sup_{\tilde{y} \in K_y} \left|\Phi\left(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{y}}\right) - \Phi\left(\tilde{\boldsymbol{x}}_2, \tilde{\boldsymbol{y}}\right)\right| \leq c_{K_y} \left|\tilde{\boldsymbol{x}}_1 - \tilde{\boldsymbol{x}}_2\right|,$$

and for every compact set $K_x \subseteq \mathcal{X}$, there exists a constant $c_{K_x} > 0$ such that for every $\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2 \in \tilde{\mathcal{Y}}$,

$$\sup_{\tilde{\boldsymbol{x}} \in K_x} \left|\Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}_1\right) - \Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}_2\right)\right| \leq c_{K_x} \left|\tilde{\boldsymbol{y}}_1 - \tilde{\boldsymbol{y}}_2\right|.$$

(ii) (Twist Condition) For every $\tilde{\boldsymbol{x}} \in \tilde{\mathcal{X}}$ *and* $\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2 \in \tilde{\mathcal{Y}}$,

$$\nabla_{(\boldsymbol{x},\boldsymbol{\varepsilon})} \Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}_1\right) = \nabla_{(\boldsymbol{x},\boldsymbol{\varepsilon})} \Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}_2\right) \implies \tilde{\boldsymbol{y}}_1 = \tilde{\boldsymbol{y}}_2,$$

and for every $\tilde{\boldsymbol{y}} \in \tilde{\mathcal{Y}}$ *and* $\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2 \in \tilde{\mathcal{X}}$,

$$\nabla_{(\boldsymbol{y},\boldsymbol{\eta})} \Phi\left(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{y}}\right) = \nabla_{(\boldsymbol{y},\boldsymbol{\eta})} \Phi\left(\tilde{\boldsymbol{x}}_2, \tilde{\boldsymbol{y}}\right) \implies \tilde{\boldsymbol{x}}_1 = \tilde{\boldsymbol{x}}_2.$$

Assumption 1 guarantees the differentiability and integrability of the production function.

Assumption 2. *(Continuous Types) The Marginal distributions* $F^u(\cdot,\cdot)$ *and* $F^d(\cdot,\cdot)$ *on* $\tilde{\mathcal{X}} = \mathbb{R}^{d_x+d_\epsilon}$ *and* $\tilde{\mathcal{Y}} = \mathbb{R}^{d_y+d_\eta}$ *are continuous.*

If Assumptions 1 and 2 hold, then there is a unique pure solution μ^* to the *primal problem*

$$\sup_{\mu \in \mathcal{M}\left(F^u, F^d\right)} \mathbb{E}_\mu \left[\Phi\left(\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{Y}}\right)\right] \tag{3}$$

such that the pure solution is characterized by two bijections, i.e., the observed and unobserved upstream matching functions, $T^u(\cdot,\cdot)$ and $S^u(\cdot,\cdot)$ or equivalently, the inverse mappings, i.e., the observed and unobserved downstream matching functions, $T^d(\cdot,\cdot)$ and $S^d(\cdot,\cdot)$. If Assumptions (1) and (2) hold, then there are π^{u*} and π^{d*} to the the *dual problem*

$$\begin{aligned} \inf_{\pi^u, \pi^d} &\left\{\mathbb{E}_{F^u}\left[\pi^u\left(\tilde{\boldsymbol{X}}\right)\right] + \mathbb{E}_{F^d}\left[\pi^d\left(\tilde{\boldsymbol{Y}}\right)\right]\right\} \\ &\text{s.t. } \pi^u\left(\tilde{\boldsymbol{x}}\right) + \pi^d\left(\tilde{\boldsymbol{y}}\right) \geq \Phi\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}\right). \end{aligned} \tag{4}$$

We note that these solutions are unique up to an additive constant and are almost-everywhere differentiable. In addition, the value of the primal solution in (3) and of the dual solutions in (4) are equal. This is known as the Monge-Kantorovich duality result. See Carlier (2003).

An important implication of this duality result is that there is a unique equilibrium matching (T^u, S^u) and unique (up to an additive constant) equilibrium profit functions $\pi^u(\cdot,\cdot)$ and $\pi^d(\cdot,\cdot)$. In other terms, we can shift $\pi^u(\cdot,\cdot)$ upward by a constant $c > 0$ and shift $\pi^d(\cdot,\cdot)$ downward by the same constant to reach new equilibrium profit functions $\tilde{\pi}^u(\cdot,\cdot)$ and $\tilde{\pi}^d(\cdot,\cdot)$ such that

$$\tilde{\pi}^u(\boldsymbol{x},\boldsymbol{\varepsilon}) = \pi^u(\boldsymbol{x},\boldsymbol{\varepsilon}) + c, \quad \tilde{\pi}^d(\boldsymbol{y},\boldsymbol{\eta}) = \pi^d(\boldsymbol{y},\boldsymbol{\eta}) - c,$$

for all $(\boldsymbol{x},\boldsymbol{\varepsilon}) \in \tilde{\mathcal{X}}$ and $(\boldsymbol{y},\boldsymbol{\eta}) \in \tilde{\mathcal{Y}}$.

With this new equilibrium result at hand, we derive the equilibrium *prices* or *transfers.* The pairwise match production can be written as the sum of the upstream and downstream valuations of the match as

$$\Phi(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) = \Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) + \Phi^d(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}),$$

where $\Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta})$ is the *upstream valuation* and $\Phi^d(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta})$ is the *downstream valuation* of the potential match between $(\boldsymbol{x},\boldsymbol{\varepsilon})$ and $(\boldsymbol{y},\boldsymbol{\eta})$. The firm's valuation of a match is the utility of being in a match before any monetary exchange or utility transfer between the two firms. The pre-transfer utility is potentially different from the firm's equilibrium profits or payoffs that can be written as functions of the equilibrium firms' valuations and the equilibrium transfer. Let $(\boldsymbol{x},\boldsymbol{\varepsilon})$-firm to be matched to $(T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}}))$-firm in a pure equilibrium matching with equilibrium payoffs $\pi^u(\boldsymbol{x},\boldsymbol{\varepsilon})$ and $\pi^d(T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}}))$. The equilibrium transfer or price in a pairwise stable equilibrium is then defined as the amount of money $p^u(\boldsymbol{x},\boldsymbol{\varepsilon})$ received by the upstream firm or $p^d(\boldsymbol{y},\boldsymbol{\eta})$ paid by the downstream firm in equilibrium. These prices/transfers defined as

$$p^u(\boldsymbol{x},\boldsymbol{\varepsilon}) \equiv \pi^u(\boldsymbol{x},\boldsymbol{\varepsilon}) - \Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon}, T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}})),$$

or equivalently for $(\boldsymbol{y},\boldsymbol{\eta})$-firm

$$p^d(\boldsymbol{y},\boldsymbol{\eta}) \equiv \Phi^d\left(T^d(\tilde{\boldsymbol{y}}), S^d(\tilde{\boldsymbol{y}}), \boldsymbol{y},\boldsymbol{\eta}\right) - \pi^d(\boldsymbol{y},\boldsymbol{\eta}).$$

In parallel, for each pairwise stable equilibrium, we define potentially several competitive equilibria by specifying the prices/transfers for matches that do not form in equilibrium. A (pure) competitive equilibrium in the continuous two-sided matching model is a pure matching characterized by (T^u, S^u) and a competitive equilibrium price function $\mathfrak{p}: \tilde{\mathcal{X}} \times \tilde{\mathcal{Y}} \mapsto \mathbb{R}$ such that for each $(\boldsymbol{x}.\boldsymbol{\varepsilon})$-firm

$$\begin{pmatrix} T^u(\boldsymbol{x},\boldsymbol{\varepsilon}) \\ S^u(\boldsymbol{x},\boldsymbol{\varepsilon}) \end{pmatrix} = \underset{(\boldsymbol{y},\boldsymbol{\eta})\in\tilde{\mathcal{Y}}}{\arg\max} \left\{\Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon}, T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}})) + \mathfrak{p}(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta})\right\},$$

and for each $(\boldsymbol{y},\boldsymbol{\eta})$-firm

$$\begin{pmatrix} T^d(\boldsymbol{y},\boldsymbol{\eta}) \\ S^d(\boldsymbol{y},\boldsymbol{\eta}) \end{pmatrix} = \arg\max_{(\boldsymbol{x},\boldsymbol{\varepsilon})\in\tilde{\mathcal{X}}} \left\{ \Phi^d\left(T^d(\tilde{\boldsymbol{y}}), S^d(\tilde{\boldsymbol{y}}), \boldsymbol{y},\boldsymbol{\eta}\right) - \mathfrak{p}(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) \right\}.$$

In other words, the equilibrium matching solves the profit maximization problem of all the firms given the competitive equilibrium price function $\mathfrak{p}(\cdot,\cdot,\cdot,\cdot)$. Consider the pairwise stable equilibrium characterized by (T^u, S^u) and $\left(\pi^u,\pi^d\right)$, then a corresponding competitive equilibrium is given by the equilibrium matching (T^u, S^u) and the equilibrium price function given by

$$\mathfrak{p}(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) = \Phi^d(\boldsymbol{x},\varepsilon,\boldsymbol{y},\eta) - \pi^d(\boldsymbol{y},\eta).$$

The competitive equilibrium matching and firms' payoffs are the same as the ones in the pairwise stable equilibrium. As an illustration, consider any two firms $\tilde{\boldsymbol{x}}$ and $\tilde{\boldsymbol{y}}$ such that $(\boldsymbol{y},\boldsymbol{\eta}) = (T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}}))$, i.e., matched in equilibrium. The competitive equilibrium price for this match is given by

$$\begin{aligned} \mathfrak{p}(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) &= \Phi^d(\boldsymbol{x},\boldsymbol{\varepsilon},T^u(\tilde{\boldsymbol{x}}),S^u(\tilde{\boldsymbol{x}})) - \pi^d s(S^u(\tilde{\boldsymbol{x}}),S^u(\tilde{\boldsymbol{x}})) \\ &= \Phi^d(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) - \pi^d(\boldsymbol{y},\boldsymbol{\eta}). \end{aligned}$$

Hence, the competitive equilibrium payoff of a $(\boldsymbol{x},\boldsymbol{\varepsilon})$-firm, denoted $\pi^{u,COMP}(\boldsymbol{x},\boldsymbol{\varepsilon})$, is given by

$$\begin{aligned} \pi^{u,COMP}(\boldsymbol{x},\boldsymbol{\varepsilon}) &= \Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) + \mathfrak{p}(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) \\ &= \Phi^u(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) + \Phi^d(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) - \pi^d(\boldsymbol{y},\boldsymbol{\eta}) \\ &= \Phi(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}) - \pi^d(\boldsymbol{y},\boldsymbol{\eta}) \\ &= \pi^u(\boldsymbol{x},\boldsymbol{\varepsilon}). \end{aligned}$$

Similarly, we can show the $(\boldsymbol{y},\boldsymbol{\eta})$-firm's competitive equilibrium profit is equal to its pairwise stable equilibrium profit. Consequently, the pairwise stable transfers coincide with the competitive equilibrium price function evaluated at the equilibrium matches. The notable difference between the stable equilibrium and the competitive equilibrium is that the equilibrium price function in the latter should be defined for *all* feasible matches, and not only those that are formed in the equilibrium. For instance, the off-equilibrium prices suggested above would make downstream firms indifferent between their equilibrium partner and any other upstream firm. The competitive equilibrium definition allows us to specify the equilibrium conditions in terms of individual profit maximization behavior for price-taking firms. Nevertheless, both notions of equilibrium imply the same model outcome—that is, the same unique equilibrium matching, equilibrium payoffs, and transfers/prices for the matches that are formed in equilibrium.

Matching Under Independence and Separability Assumptions

We now discuss the equilibrium under more restrictive assumptions such as independence of the unobserved and observed types and additive separability of the production function. We state these assumptions below. Let $F^u(\cdot,\cdot)$ and $F^d(\cdot,\cdot)$ denote the marginal distributions of observed and unobserved types of the upstream and downstream firms, respectively. The induced marginal and conditional distributions of the firms' types are denoted by the subscripts.

Assumption 3. *(Independence) The upstream unobserved types $\boldsymbol{\varepsilon}$ are independently distributed from the observed types $\boldsymbol{X}$, namely*

$$F^u\left(\boldsymbol{x},\boldsymbol{\varepsilon}\right)=F^u_{\varepsilon|x}\left(\boldsymbol{\varepsilon}|\boldsymbol{X}=\boldsymbol{x}\right)F^u_x\left(\boldsymbol{x}\right)=F^u_\varepsilon\left(\boldsymbol{\varepsilon}\right)F^u_x\left(\boldsymbol{x}\right).$$

Similarly, the downstream unobserved types $\boldsymbol{\eta}$ are independently distributed from the observed types $\boldsymbol{Y}$, namely

$$F^d\left(\boldsymbol{y},\boldsymbol{\eta}\right)=F^d_{\eta|y}\left(\boldsymbol{\eta}|\boldsymbol{Y}=\boldsymbol{y}\right)F^d_y\left(\boldsymbol{y}\right)=F^d_\eta\left(\boldsymbol{\eta}\right)F^d_y\left(\boldsymbol{y}\right).$$

Assumption 4. *(Separability) The pairwise match production function is additively separable in the observed and unobserved types, namely potential match production of a match between a $(\boldsymbol{x},\boldsymbol{\varepsilon})$-firm and a $(\boldsymbol{y},\boldsymbol{\eta})$-firm is given by $\Phi\left(\boldsymbol{x},\boldsymbol{\varepsilon},\boldsymbol{y},\boldsymbol{\eta}\right)=\zeta\left(\boldsymbol{x},\boldsymbol{y}\right)+\Xi\left(\boldsymbol{\varepsilon},\boldsymbol{\eta}\right)$, where $\zeta\left(\cdot,\cdot\right)$ and $\Xi\left(\cdot,\cdot\right)$ satisfy Assumption 1.*

Next, we establish a theorem that derives the pure equilibrium matching under these assumptions.

Theorem 1. *Let Assumptions 1–4 hold. The unique equilibrium matching is characterized by $\left(\tilde{T}^u\left(\cdot\right),\tilde{S}^u\left(\cdot\right)\right)$ where $\tilde{T}^u:\mathbb{R}^{d_x}\mapsto\mathbb{R}^{d_y}$ and $\tilde{S}^u:\mathbb{R}^{d_\epsilon}\mapsto\mathbb{R}^{d_\eta}$—that is, for each $(\boldsymbol{x},\boldsymbol{\varepsilon})$-firm the observed types of the equilibrium partner are only a function of the observed types $\boldsymbol{x}$, and the unobserved types of the equilibrium partner are only a function of the unobserved types $\boldsymbol{\varepsilon}$.*

The uniqueness of the equilibrium matching is guaranteed by the Monge-Kantorovich theorem under Assumptions 1 and 2. The proof of Theorem 1 relies on rewriting the primal problem (3) as two separate maximization problems under the independence assumption. This means that the matchings of the observed and unobserved types are independent from each other. Theorem 1 implies that upstream firms characterized by a specific observed type $\boldsymbol{x}$ in a dataset of equilibrium matching generated by a model under Assumptions 1–4 are matched to downstream firms with a unique observed type $\boldsymbol{y}$. We say that the observed matching is deterministic. For instance, if $d_x = d_y = 1$, i.e. the case of scalar observed types, the equilibrium matching observed in the data should be a line in the $x - y$ space. This model implication is rarely satisfied in practice where we expect to observe firms with similar observed types to be matched to firms that do not have the same observed types. Therefore, the model is misspecified. This is not due to lack of model flexibility in terms of the unobserved heterogeneity as we can include complex unobserved heterogeneity structure in $\Xi(\cdot,\cdot)$, but rather the property of the equilibrium matching in a continuous market under Assumptions (1)–(3).

3 Identification under Profit Data Availability

In this section, we consider the identification and estimation of the model under different model assumptions and data schemes. The model primitives consist of the match production function and the marginal distributions of the upstream and downstream characteristics, i.e. $\Phi(\cdot,\cdot,\cdot,\cdot)$, $F^u(\cdot,\cdot)$, and $F^d(\cdot,\cdot)$. In terms of data schemes, we consider the cases where the analyst has access to either individual profits for all the firms or only firms on one side of the market, or match transfers/prices for all the matches that are formed in equilibrium. Let the upstream firms in the observed sample be indexed by $i = 1, 2, \ldots, N$. Denote by $(\boldsymbol{x}_i, \boldsymbol{y}_i^*)$ the observed match between the upstream firm i and its downstream equilibrium partner, where $\boldsymbol{x}_i$ is a vector of the observed characteristics of the upstream firm i. Similarly, $\boldsymbol{y}_i^*$ is the vector of observed characteristics of the downstream firm that is matched to the upstream firm i. The upstream and downstream profits of the match indexed by i are denoted by π_i^u and π_i^d, respectively. Likewise, the transfer/price associated with observation i's match is denoted by p_i.

Data Scheme 1. *The data include the equilibrium matching and individual profits of all firms from one large market. The observed sample is* $\left\{(\boldsymbol{x}_i, \boldsymbol{y}_i^*), \pi_i^u, \pi_i^d\right\}_{i=1}^N$.

Data Scheme 2. *The data include the equilibrium matching and individual profits of either the upstream side or the downstream side, but not both, from one large market. The observed sample is* $\{(\boldsymbol{x}_i, \boldsymbol{y}_i^*), \pi_i^u\}_{i=1}^N$ *or* $\left\{(\boldsymbol{x}_i, \boldsymbol{y}_i^*), \pi_i^d\right\}_{i=1}^N$.

Data Scheme 3. *The data include the equilibrium matching and match transfer for all the matches that formed in equilibrium from one large market. The observed sample is* $\{(\boldsymbol{x}_i, \boldsymbol{y}_i^*), p_i\}_{i=1}^N$.

3.1 Identification with Profit Data under Monotonicity and Independence

We first consider the model under Data Scheme 1 where the analyst has access to the matching and individual profits of all firm from one large market. We also discuss how data from multiple markets can be used for identification purpose.

The equilibrium matching, either competitive or pairwise, can be thought as the result of profit maximization of all the firms. At equilibrium, an upstream firm characterized by $(\boldsymbol{x}, \varepsilon)$ chooses the downstream firm to maximize its equilibrium profit. The upstream firm chooses $(\boldsymbol{y}, \boldsymbol{\eta})$ given the downstream equilibrium profit function $\pi^u(\boldsymbol{x}, \varepsilon)$chooses $(\boldsymbol{y}, \eta)$ so as to maximize the potential match production net of the equilibrium profit share of the downstream firm, given by $\Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) - \pi^d(\boldsymbol{y}, \eta)$. The profit maximization implies given the downstream equilibrium function $\pi^d(\cdot, \cdot)$, the upstream profit function at every point $(\boldsymbol{x}, \varepsilon) \in \mathcal{X}$ is given by

$$\pi^u(\boldsymbol{x}, \varepsilon) = \max_{(\boldsymbol{y}, \eta) \in \mathcal{Y}} \left\{\Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) - \pi^d(\boldsymbol{y}, \eta)\right\}. \tag{5}$$

Suppose that for some $(\boldsymbol{x}, \varepsilon)$ matched to $(\boldsymbol{y}, \eta)$ at equilibrium, (2) does not hold. Then, there should be another $(\boldsymbol{y}', \eta') \neq (\boldsymbol{y}, \eta)$ such that

$$\Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}', \eta') - \pi^d(\boldsymbol{y}', \eta') > \Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) - \pi^d(\boldsymbol{y}, \eta).$$

By (2), we can replace $\Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta)$ by $\pi^u(\boldsymbol{x}, \varepsilon) + \pi^d(\boldsymbol{y}, \eta)$. Thus, we have

$$\pi^u(\boldsymbol{x}, \varepsilon) + \pi^d(\boldsymbol{y}', \eta') < \Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}', \eta'),$$

which violates the condition for pairwise stable equilibrium in (1). Similarly, given the equilibrium upstream profit function, the equilibrium downstream profit function should satisfy

$$\pi^d(\boldsymbol{y}, \eta) = \max_{(\boldsymbol{x}, \varepsilon) \in \mathcal{X}} \left\{\Phi(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) - \pi^u(\boldsymbol{x}, \varepsilon)\right\},$$

at all points $(\boldsymbol{y}, \eta) \in \tilde{\mathcal{Y}}$—in other words, π^d should satisfy the profit maximization of the downstream firm.

Let $\pi^u_{x_k}$ and $\pi^d_{y_k}$ denote the partial derivative of the upstream and downstream profit functions with respect to the argument corresponding to the k'th observable characteristic. Similarly, let π^u_ε and π^d_η denote the partial derivatives with respect to the argument corresponding to the scalar unobserved characteristic. [1]

The following assumption imposes monotonicity restrictions on the structural match production function which in turn implies monotonicity of the equilibrium profit functions—which is the equilibrium outcome of the model.

Assumption 5. *(Monotonicity) The match production function $\Phi(\cdot,\cdot,\cdot,\cdot)$ is strictly monotonic in the scalar unobservable characteristics ε and η. That is,*

$$\Phi_\varepsilon(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) > 0 \quad \textit{and} \quad \Phi_\eta(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) > 0,$$

for all $(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{Y}}$.

The next lemma is a simple application of the envelope theorem under the monotonicity assumption. [2]

Lemma 1. *Let Assumptions 1,2, and 5 hold. The equilibrium profit functions of the upstream and downstream firms are monotonic in the unobservable characteristics, that is*

$$\pi^u_\varepsilon(\boldsymbol{x}, \varepsilon) > 0 \quad \textit{and} \quad \pi^d_\eta(\boldsymbol{y}, \eta) > 0.$$

The envelope theorem says that to find the partial derivatives of the value function π^u with respect to its parameters parameters $\boldsymbol{x}$ and ε in (5), we only need to focus on direct effects of the parameters—i.e., indirect effects through the maximizers $T^u(\boldsymbol{x}, \varepsilon)$ and $S^u(\boldsymbol{x}, \varepsilon)$ can be ignored. The partial derivative of π^u with respect to ε is given by

[1] In the case of a vector of unobservables, we include a subscript to specify the components of the vector.

[2] e.g., Simon and Blume (1994, pp. 453–455)

$$
\begin{aligned}
\frac{\partial \pi^u(\boldsymbol{x},\varepsilon)}{\partial \varepsilon} =& \frac{\partial\left[\Phi\left(\boldsymbol{x},\varepsilon,T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right)\right)-\pi^d\left(T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right)\right)\right]}{\partial \varepsilon} \\
=&\Phi_\varepsilon\left(\boldsymbol{x},\varepsilon,T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right)\right) \\
&+\sum_{k=1}^{d_y}\frac{\partial T_k^u\left(\tilde{\boldsymbol{x}}\right)}{\partial\varepsilon}\left[\Phi_{y_k}(\boldsymbol{x},\varepsilon,T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right))-\pi^d_{y_k}(T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right))\right] \qquad (6)\\
&+\frac{\partial S^u(\tilde{\boldsymbol{x}})}{\partial\varepsilon}\left[\Phi_\eta(\boldsymbol{x},\varepsilon,T^u(\tilde{\boldsymbol{x}}),S^u(\tilde{\boldsymbol{x}}))-\pi^d_\eta\left(T^u(\tilde{\boldsymbol{x}}),S^u(\tilde{\boldsymbol{x}})\right)\right] \\
=&\Phi_\varepsilon\left(\boldsymbol{x},\varepsilon,T^u\left(\tilde{\boldsymbol{x}}\right),S^u\left(\tilde{\boldsymbol{x}}\right)\right).
\end{aligned}
$$

The third equality is implied by the envelope theorem. The indirect effects through the optimizers $\partial/\partial\varepsilon T_k^u\left(\boldsymbol{x},\varepsilon\right)$ and $\partial/\partial\varepsilon S^u\left(\boldsymbol{x},\varepsilon\right)$ can be ignored. Therefore, Assumption (5) implies that the equilibrium profits are monotonic in ε. A similar argument ensures the monotonicity in the downstream unobservable η.

Let $\alpha_\varepsilon \equiv F_\varepsilon^u\left(\varepsilon\right)$ and $\alpha_\eta \equiv F_\eta^d\left(\eta\right)$. By construction, the random variables α_ε and α_η are both distributed uniformly on $[0,1]$. Let $\bar{T}^u\left(\boldsymbol{x},\alpha_\varepsilon\right) \equiv T^u\left(\boldsymbol{x},F_\varepsilon^{u-1}\left(\alpha_\varepsilon\right)\right)$ and $\bar{S}^u\left(\boldsymbol{x},\alpha_\varepsilon\right) \equiv S^u\left(\boldsymbol{x},F_\varepsilon^{u-1}\left(\alpha_\varepsilon\right)\right)$ be the observed and unobserved equilibrium matching functions evaluated at observed characteristics $\boldsymbol{x}$ and the α_ε-quantile of the unobserved distribution F_ε^u. Similarly, $\bar{\pi}^u\left(\boldsymbol{x},\alpha_\varepsilon\right) \equiv \pi^u\left(\boldsymbol{x},F_{\alpha_\varepsilon}^{u\ -1}\left(\alpha_\varepsilon\right)\right)$, $\bar{\pi}^d\left(\boldsymbol{y},\alpha_\eta\right) \equiv \pi^d\left(\boldsymbol{y},F_\eta^{d-1}\left(\alpha_\eta\right)\right)$, and $\bar{\Phi}\left(\boldsymbol{x},\alpha_\varepsilon,\boldsymbol{y},\alpha_\eta\right) \equiv \Phi(\boldsymbol{x},F_\varepsilon^{u-1}\left(\alpha_\varepsilon\right)\boldsymbol{y},F_\eta^{d-1}\left(\alpha_\varepsilon\right))$.

The next result builds on Matzkin (2003) who studies the identification of nonseparable functions under independence and monotonicity in an unobserved scalar. The equilibrium profits are nonseparable functions of the observed vectors and an unobserved scalar, as for instance, the upstream profits can be written as $\pi^u(\boldsymbol{x},\varepsilon) = \Phi(\boldsymbol{x},\varepsilon,T^u\left(\boldsymbol{x},\varepsilon\right),S^u(\boldsymbol{y},\eta)) \equiv m(\boldsymbol{x},\varepsilon)$.

Theorem 2. *Let Assumptions 1,2,3, and 5; and Data Scheme 1 hold. The match production function evaluated at quantiles of the unobserved distributions*

$$\bar{\Phi}\left(\boldsymbol{x},\alpha_\varepsilon,\boldsymbol{y},\alpha_\eta\right)$$

is identified on the equilibrium path $\left(\boldsymbol{x},\alpha_\varepsilon,\bar{T}^u\left(\boldsymbol{x},\alpha_\varepsilon\right),\bar{S}^u\left(\boldsymbol{x},\alpha_\varepsilon\right)\right)$ *for all* $\boldsymbol{x}$ *and all* $\alpha_\varepsilon \in [0,1]$.

The proof is provided in the appendix. The proof shows that under independence and monotonicity assumptions, for each upstream firm the quantile α_ε and for each downstream firm the quantile α_η of their unobserved characteristics are identified as the distribution of the upstream and downstream profits conditional on

$\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. The upstream profit data identify $\bar{\pi}^u(\boldsymbol{x}, \alpha_\varepsilon)$ as the profit of the upstream firm with observed characteristics $\boldsymbol{x}$ and the unobservable α_ε, which is already pinned down for each upstream firm. Similarly, for $\bar{\pi}^d(\boldsymbol{y}, \alpha_\eta)$. The matching data identifies $\bar{T}^u(\boldsymbol{x}, \alpha_\varepsilon)$ as the observed characteristic $\boldsymbol{y}$ of the equilibrium partner of $(\boldsymbol{x}, \alpha_\varepsilon)$, and $\bar{S}^u(\boldsymbol{x}, \alpha_\varepsilon)$ is identified as the unobserved characteristic α_η of the equilibrium partner of $(\boldsymbol{x}, \alpha_\varepsilon)$. The last step of the proof uses (2) to identify$\bar{\Phi}\left(\boldsymbol{x}, \bar{\varepsilon}, \bar{T}^u(\boldsymbol{x}, \bar{\varepsilon}), \bar{S}^u(\boldsymbol{x}, \bar{\varepsilon})\right)$ for all $\boldsymbol{x}$ and $\bar{\varepsilon} \in [0,1]$ as the sum of the matched upstream and downstream firms, namely

$$\bar{\Phi}\left(\boldsymbol{x}, \bar{\varepsilon}, \bar{T}^u(\boldsymbol{x}, \bar{\varepsilon}), \bar{S}^u(\boldsymbol{x}, \bar{\varepsilon})\right) = \pi^u(\boldsymbol{x}, \bar{\varepsilon}) + \pi^d\left(\bar{T}^u(\boldsymbol{x}, \bar{\varepsilon}), \bar{S}^u(\boldsymbol{x}, \bar{\varepsilon})\right).$$

It is worth noting that Theorem 2 under Data Scheme 1 does not identify the normalized match production function $\bar{\Phi}(\cdot,\cdot,\cdot,\cdot)$ on all of its domain. For instance, an upstream firm with observed characteristics $\boldsymbol{x}$ receiving the median profit conditional on $\boldsymbol{X} = \boldsymbol{x}$, i.e., $\bar{\pi}^u\left(\boldsymbol{x}, \frac{1}{2}\right)$, is matched to only one downstream firm in equilibrium that is characterized by the observed characteristics $\boldsymbol{y} = \bar{T}^u\left(\boldsymbol{x}, \frac{1}{2}\right)$ and is receiving the $\bar{S}^u\left(\boldsymbol{x}, \frac{1}{2}\right)$-quantile of downstream profits conditional on $\boldsymbol{Y} = \boldsymbol{y}$, namely $\bar{\pi}^d(\boldsymbol{y}, \bar{S}^u(\boldsymbol{x}, \frac{1}{2}))$. In other words, data from one market allow us to identify $\bar{\Phi}(\cdot,\cdot,\cdot,\cdot)$ at the point $\left(\boldsymbol{x}, \frac{1}{2}, \bar{T}^u\left(\boldsymbol{x}, \frac{1}{2}\right), \bar{S}^u\left(\boldsymbol{x}, \frac{1}{2}\right)\right)$ but not off equilibrium points $(\boldsymbol{x}, \alpha_\varepsilon, \boldsymbol{y}, \alpha_\eta)$ where $\boldsymbol{y} \neq \bar{T}^u(\boldsymbol{x}, \alpha_\varepsilon)$ and/or $\bar{\eta} \neq \bar{S}^u(\boldsymbol{x}, \alpha_\eta)$.

We propose two solutions to the nonidentification of the match production function at the off-equilibrium points. A first solution is to obtain additional data from other markets with different distributions of characteristics, but sharing the same structural match production function. For each point $(\boldsymbol{x}, \alpha_\varepsilon, \boldsymbol{y}, \alpha_\eta)$, the match production function is identified from a market where the point is an equilibrium matching in that market. Alternatively, we can consider a semiparametric model where the match production function is approximated by a flexible function of the characteristics such as using a high degree polynomial in the observed and unobserved characteristics. This strategy has been used in the literature with data on matching to estimate features of the match production function such as complementarities, i.e., second derivatives, see e.g. Chiappori et al. (2022).

Our next result focuses on the implication of the separability assumption of the match production function and the result in Theorem 1 on the identification result of Theorem 2.

Corollary 1. *Let Assumptions 1–5 and Data Scheme 1 hold. The observed match production function* $\zeta(\cdot,\cdot)$ *is identified on* $(\boldsymbol{x}, T^u(\boldsymbol{x}))$ *for all* $\boldsymbol{x}$*, and the unobserved* *match production function of the normalized unobserved characteristics* $\bar{\Xi}(\cdot,\cdot)$ *on* $(\alpha_\varepsilon, S^u(\alpha_\varepsilon))$ *for all* $\alpha_\varepsilon \in [0,1]$.

Under the assumptions of Corollary 1, the observed matching $(\boldsymbol{x}, T^u(\boldsymbol{x}))$ is just a line—similarly for the unobserved matching $(\alpha_\varepsilon, S^u(\alpha_\varepsilon))$. In other words, the observed production is identified only on one line.

Pairwise Stability Bounds

The data from one market without assuming a parametric functional form is still informative about the production of the matches that are not formed in equilibrium. We rewrite the pairwise stability condition in (4)

$$\bar{\pi}^u(\boldsymbol{x}, \alpha_\varepsilon) + \bar{\pi}^d(\boldsymbol{y}, \alpha_\eta) \geq \bar{\Phi}(\boldsymbol{x}, \alpha_\varepsilon, \boldsymbol{y}, \alpha_\eta), \tag{7}$$

which holds for all upstream $(\boldsymbol{x}, \alpha_\varepsilon)$ and downstream $(\boldsymbol{y}, \alpha_\eta)$ even if they are not matched to each other in equilibrium. The left-hand side of (7) are the profits that the firms receive from their equilibrium matches, while the right-hand side is the potential production of the match between $(\boldsymbol{x}, \alpha_\varepsilon)$ and $(\boldsymbol{y}, \alpha_\eta)$ which is not formed in equilibrium. This means that without further assumptions on the production function, the model provides an upper bound for the production of all those matches that are not formed in equilibrium—i.e., the points at which the production function is not identified without further data or parametric assumptions. When the production function is approximated by a flexible parametric functional form, we can use the additional information provided by these upper bounds in the estimation.

Next, we construct an example with a quadratic specification of the production function.

Example 1. *(Semiparametric Quadratic production)* Let x, ε, y, η be scalar random variables and Assumptions 1–3 hold. Let the total production be

$$\Phi(x, \varepsilon, y, \eta) \equiv \beta_{xx}x^2 + \beta_{xy}xy + \beta_{yy}y^2 + \beta_{x\eta}x\eta + \beta_{y\varepsilon}y\varepsilon + \varepsilon\eta, \tag{8}$$

so that

$$\Phi_2(x, \varepsilon, y, \eta) = \beta_{y\varepsilon}y + \eta \text{ and } \Phi_4(x, \varepsilon, y, \eta) = \beta_{x\eta}x + \varepsilon.$$

Further assume $x, y, \varepsilon, \eta > 0$ and $\beta_{y\varepsilon}, \beta_{x\eta} > 0$ so that $\pi^u(x, \varepsilon)$ and $\pi^d(y, \eta)$ are monotone in ε and η, respectively. For each matched pair reported in the sample, which is indexed by i, we observe $\left(x_i, y_i, \pi_i^u, \pi_i^d\right)$. Let ε_i,η_i denote the latent unobserved characteristics. By definition of equilibrium, we know that

$$
\begin{aligned}
\pi_i^u &= \pi^u(x_i, \varepsilon_i),\ \pi_i^d = \pi^d(y_i, \eta_i), \\
y_i &= T^u(x_i, \varepsilon_i),\ \eta_i = S^u(x_i, \varepsilon_i), \\
x_i &= T^d(y_i, \eta_i),\ \varepsilon_i = S^d(y_i, \eta_i).
\end{aligned}
$$

Given the profit monotonicity and observation of $\left(x_i, y_i, \pi_i^u, \pi_i^d\right)$ in the data, we can back out the quantiles of ε and η in each matched pair in the data. Specifically, for each matched pair i, we cannot directly pin down the realized values of ε_i and η_i, but we can calculate quantiles $\alpha_{\varepsilon i}$ and $\alpha_{\eta i}$ such that

$$
\begin{aligned}
\Pr(\pi^u &\leq \pi_i^u | X = x_i) = \Pr(\varepsilon \leq \varepsilon_i) = \alpha_{\varepsilon i}, \\
\Pr(\pi^d &\leq \pi_i^d | Y = y_i) = \Pr(\eta \leq \eta_i) = \alpha_{\eta i}.
\end{aligned}
$$

An application of the envelop theorem in the dual of the Kantorovich problem implies that

$$
\begin{aligned}
\pi_1^u(x, \varepsilon) &= \Phi_1(x, \varepsilon, T^u(x, \varepsilon), S^u(x, \varepsilon)) = 2\beta_{xx}x + \beta_{xy}T^u(x, \varepsilon) + \beta_{x\eta}S^u(x, \varepsilon), \\
\pi_1^d(y, \eta) &= \Phi_3(T^d(y, \eta), S^d(y, \eta), y, \eta) = 2\beta_{yy}y + \beta_{xy}T^d(y, \eta) + \beta_{y\varepsilon}
\end{aligned}
$$

Following the independence assumption, we can identify $\pi_1^u\left(x_i, \varepsilon_i\right)$ as $\nabla_x Q^u\left(\alpha_{\varepsilon i}|x_i\right)$ for each matched pair i in the data, where $Q^u\left(\alpha|x_i\right)$ denotes the α-th conditional quantile of upstream profit π^u conditional on $X = x_i$ and $\nabla_x Q^u\left(\alpha_i^\varepsilon|x_i\right)$ is the partial derivative with respect to x. This gives the following equations

$$
\begin{aligned}
2\beta_{xx}x_i + \beta_{xy}y_i + \beta_{x\eta}\eta_i &= \nabla_x Q^u(\alpha_{\varepsilon i}|x_i) \equiv u_i, \\
2\beta_{yy}y_i + \beta_{xy}x_i + \beta_{y\varepsilon}\varepsilon_i &= \nabla_y Q^d(\alpha_{\eta i}|y_i) \equiv d_i.
\end{aligned}
$$

Consider any two observations of matched pairs i, j such that $\alpha_{\eta i} = \alpha_{\eta j}$ (or equivalently $\eta_i = \eta_j$). We take the difference to get

$$
\begin{aligned}
2\beta_{xx}(x_i - x_j) + \beta_{xy}(y_i - y_j) &= u_i - u_j, \\
2\beta_{yy}(y_i - y_j) + \beta_{xy}(x_i - x_j) &= d_i - d_j.
\end{aligned}
$$

Assuming the paired random vector $(X_i - X_j, Y_i - Y_j)$ is not contained in a linear subspace conditional on $\eta_i = \eta_j$, we can (over-)identify β_{xx} and β_{xy} using the first implication. By an analogous argument, we can (over-)identify β_{xx} and β_{xy} using the second as well. This in turn implies we can identify

$$\beta_{x\eta}\eta_i = u_i - 2\beta_{xx}x_i - \beta_{xy}y_i, \tag{9}$$

$$\beta_{y\varepsilon}\varepsilon_i = d_i - 2\beta_{yy}y_i - \beta_{xy}x_i, \tag{10}$$

for each matched pair i in the sample.

It remains to recover $\beta_{x\eta}$, $\beta_{y\varepsilon}$, and the marginal distribution of ε and η. To do so, recall that we can identify $\Phi(\cdot,\cdot,\cdot,\cdot)$ over the equilibrium matching support $(x,\varepsilon,T^u(x,\varepsilon),S^u(x,\varepsilon))$. That is we can recover

$$\begin{aligned}\Phi(x,\varepsilon,T^u(x,\varepsilon),S^u(x,\varepsilon)) = \beta_{xx}x^2 + \beta_{xy}xT^u(x,\varepsilon) + \beta_{yy}T^u(x,\varepsilon)^2 + \beta_{x\eta}xS^u(x,\varepsilon)\\ + \beta_{y\varepsilon}T^u(x,\varepsilon)\varepsilon + \varepsilon S^u(x,\varepsilon),\end{aligned}$$

for all (x,ε). In terms of observed information from the sample, this means we can observe total profits for each matched pair i, their respective observed characteristics (x_i,y_i), but not the latent (ε_i,η_i) in the following equation,

$$\Phi_i = \beta_{xx}x_i^2 + \beta_{xy}x_iy_i + \beta_{yy}y_i^2 + \beta_{x\eta}x_i\eta_i + \beta_{y\varepsilon}y_i\varepsilon_i + \varepsilon_i\eta_i. \tag{11}$$

However, as explained above, we do know the corresponding percentile values for ε_i and η_i, i.e. $(\alpha_{\varepsilon i},\alpha_{\eta i})$, for each i. By substituting (9) and (10) into (11), we can express $\varepsilon_i\eta_i$ in terms of identifiable quantities.

Next, note that we need at least one scale normalization, in addition to normalizing the coefficient in from of $\varepsilon_i\eta_i$ to one, in the distribution of unobserved characteristics ε or η. We normalize the median of the marginal distribution of η to a known value η_0. Then, we can recover $\mathcal{E}_0 \equiv \{\varepsilon_i : \exists x_i \text{ s.t. } S^u(x_i,\varepsilon_i) = \eta_0\}$ using the knowledge of $\varepsilon_i\eta_i$ on the equilibrium match support. Specifically, for an observation with $\alpha_{\eta i} = 0.5$, we can identify ε_i simply as $\varepsilon_i\eta_i$, which is identified above. More importantly, we already know this realized value of ε_i corresponds to the $\alpha_{\varepsilon i}$-th percentile in the marginal distribution of ε_i, because $\alpha_{\varepsilon i}$ is identified earlier using monotonicity of $\pi^u(x,\varepsilon)$ in ε. Next, for each $\varepsilon_0 \in \mathcal{E}_0$ we can recover $\mathcal{H}_0 \equiv \{\eta_i : \exists y_i \text{ s.t. } S^d(y_i,\eta_i) = \varepsilon_0\}$, again using knowledge of $\varepsilon_i\eta_i$ on equilibrium support. Specifically, for an observed matched pair i with $\varepsilon_i = \varepsilon_0 \in \mathcal{E}_0$, we identify η_i as $\varepsilon_i\eta_i/\varepsilon_0$, where the product $\varepsilon_i\eta_i$ identified as above. Again, we know η_i is the $\alpha_{\eta i}$-th percentile of the marginal distribution of η_i, thanks to previous identification of α_i^η. It then follows immediately that both $\beta_{x\eta},\beta_{y\varepsilon}$ are identified from (9) and

(10). Furthermore, iterating this process to expand the sets to $\mathcal{E}_0, \mathcal{H}_0, \mathcal{E}_1, \mathcal{H}_1, \ldots$ identifies more percentiles ε_i and η_i on their marginal support.

3.2 Identification with Profit Data under Separability

In this section, we consider the identification when the restrictive assumption of monotonicity, i.e. Assumption 5, and the implicit assumption of unobservable scalar in Theorem 2 are relaxed. Without these assumptions, the conditional quantiles of the profit data do not correspond to the quantiles of the unobserved variables distributions. The next result relaxes the monotonicity and unobserved scalar assumptions, but imposes the separability assumption, i.e. Assumption 4, to restore identification.

Theorem 3. *Let Assumptions 1,2, and 4, and Data Scheme 2 hold. The mean of the partial derivative of observed production function* $\mathbb{E}\left[\zeta_{x_k}\left(\boldsymbol{x}, T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)\right)|\boldsymbol{X} = \boldsymbol{x}\right]$, $k = 1, \ldots, d_x$ *is identified for all* $\boldsymbol{x} \in \mathcal{X}$.

The proof is as follows. The $(\boldsymbol{x}, \boldsymbol{\varepsilon})$-firm's profit maximization problem is given by

$$\pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) = \max_{(\boldsymbol{y}, \boldsymbol{\eta}) \in \mathcal{Y}} \left\{\zeta\left(\boldsymbol{x}, \boldsymbol{y}\right) + \Xi\left(\boldsymbol{\varepsilon}, \boldsymbol{\eta}\right) - \pi^d\left(\boldsymbol{y}, \eta\right)\right\}.$$

The envelope theorem implies that the partial derivative of the upstream profit with respect to x_k evaluated at the equilibrium is given by

$$\begin{aligned} \frac{\partial \pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)}{\partial x_k} &= \left.\frac{\partial\left[\zeta\left(\boldsymbol{x}, \boldsymbol{y}\right) + \xi\left(\boldsymbol{\varepsilon}, \boldsymbol{\eta}\right) - \pi^d\left(\boldsymbol{y}, \boldsymbol{\eta}\right)\right]}{\partial x_k}\right|_{(\boldsymbol{y}, \boldsymbol{\eta}) = (T^u(\tilde{\boldsymbol{x}}), S^u(\tilde{\boldsymbol{x}}))} \\ &= \zeta_{x_k}\left(\boldsymbol{x}, T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)\right). \end{aligned} \tag{12}$$

Taking expectation of both sides of (12) with respect to the conditional distribution of $\boldsymbol{\varepsilon}$ gives

$$\mathbb{E}_{F^u_{\varepsilon|x}}\left[\frac{\partial}{\partial x_k}\pi^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right)|\boldsymbol{X} = \boldsymbol{x}\right] = \mathbb{E}_{F^u_{\varepsilon|x}}\left[\zeta_{x_k}\left(\boldsymbol{X}, T^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right)\right)|\boldsymbol{X} = \boldsymbol{x}\right].$$

The left-hand side can be written as

$$\mathbb{E}_{F^u_{\boldsymbol{\epsilon}|\boldsymbol{X}}}\left[\frac{\partial \pi^u\left(\boldsymbol{X},\boldsymbol{\varepsilon}\right)}{\partial x_k}|\boldsymbol{X}=\boldsymbol{x}\right] = \int_{\mathcal{E}} \frac{\partial \pi^u\left(\boldsymbol{x},\boldsymbol{\varepsilon}\right) f^u_{\boldsymbol{\epsilon}|\boldsymbol{X}}\left(\boldsymbol{\varepsilon}|\boldsymbol{x}\right)\, d\boldsymbol{\varepsilon}}{\partial x_k}$$

$$= \frac{\partial \int_{\varepsilon\in\mathcal{E}} \pi^u\left(\boldsymbol{x},\boldsymbol{\varepsilon}\right) f^u_{\boldsymbol{\epsilon}|\boldsymbol{X}}\left(\boldsymbol{\varepsilon}|\boldsymbol{x}\right)\, d\boldsymbol{\varepsilon}}{\partial x_k}$$

$$= \frac{\partial\, \mathbb{E}\left[\pi^u\left(\boldsymbol{X},\boldsymbol{\varepsilon}\right)|\boldsymbol{X}=\boldsymbol{x}\right]}{\partial x_k}.$$

The last equality is using the Leibniz integral rule to change the order of integration and differentiation. Thus, we obtain

$$\frac{\partial\, \mathbb{E}\left[\pi^u\left(\boldsymbol{X},\boldsymbol{\varepsilon}\right)|\boldsymbol{X}=\boldsymbol{x}\right]}{\partial x_k} = \mathbb{E}_{F^u_{\varepsilon|x}}\left[\zeta_{x_k}\left(\boldsymbol{X},T^u\left(\boldsymbol{X},\boldsymbol{\varepsilon}\right)\right)|\boldsymbol{X}=\boldsymbol{x}\right], \tag{13}$$

where the left-hand side is identified from the data on upstream equilibrium profits.

The next example illustrates an application of Theorem 3 in a semi-parametric setting where the observed production function is linear in the parameters.

Example 2. We consider the match production $\Phi\left(x,y,\boldsymbol{\varepsilon},\boldsymbol{\eta}\right) = \beta_{xx}x^2 + \beta_{xy}xy + \beta_{yy}y^2 + \Xi\left(\boldsymbol{\varepsilon},\boldsymbol{\eta}\right)$. For each $x\in\mathbb{R}$, the expectation of the partial derivatives

$$\mathbb{E}\left[2\beta_{xx}xT^u\left(X,\boldsymbol{\varepsilon}\right) + \beta_{xy}T^u\left(X,\boldsymbol{\varepsilon}\right)|X=x\right]$$
$$= 2\beta_{xx}x\,\mathbb{E}\left[T^u\left(X,\boldsymbol{\varepsilon}\right)|X=x\right] + \beta_{xy}\,\mathbb{E}\left[T^u\left(X,\varepsilon\right)|X=x\right]$$

is identified using data on upstream firms' profits. The expectation term on the right-hand side, i.e. $\mathbb{E}\left[T^u\left(x,\boldsymbol{\varepsilon}\right)|X=x\right]$, is also identified using data on equilibrium matching as the mean of the observed characteristic of the matching partner of x-firms. We define $\bar{y}^*\left(x\right)\equiv\mathbb{E}\left[T^u\left(X,\boldsymbol{\varepsilon}\right)|X=x\right]$. Let $x_1,x_2\in\mathbb{R}$ and $x_1\neq x_2$. Then, we have

$$\frac{\partial\, \mathbb{E}\left[\pi^u\left(X,\boldsymbol{\varepsilon}\right)|X=x_1\right]}{\partial x} = \bar{y}^*\left(x_1\right)\left(2\beta_{xx}x_1+\beta_{xy}\right),$$
$$\frac{\partial\, \mathbb{E}\left[\pi^u\left(X,\boldsymbol{\varepsilon}\right)|X=x_2\right]}{\partial x} = \bar{y}^*\left(x_2\right)\left(2\beta_{xx}x_2+\beta_{xy}\right).$$

It follows that β_{xx} and β_{xy} are identified as

$$\beta_{xx} = \frac{\frac{\partial\,\mathbb{E}\left[\pi^u(X,\boldsymbol{\varepsilon})|X=x_2\right]}{\partial x}\bar{y}^*\left(x_1\right) - \frac{\partial\,\mathbb{E}\left[\pi^u(X,\boldsymbol{\varepsilon})|X=x_1\right]}{\partial x}\bar{y}^*\left(x_2\right)}{2\left(x_1-x_2\right)y_1y_2},$$

and

$$\beta_{xy} = -\frac{\frac{\partial \mathbb{E}[\pi^u(X,\boldsymbol{\varepsilon})|X=x_2]}{\partial x} x_1 \bar{y}^*(x_1) + \frac{\partial \mathbb{E}[\pi^u(X,\boldsymbol{\varepsilon})|X=x_1]}{\partial x} x_2 \bar{y}^*(x_2)}{(x_1 - x_2) y_1 y_2}.$$

Similarly, the downstream profit data identify β_{yy} as

$$\beta_{yy} = \frac{\frac{\partial \mathbb{E}\left[\pi^d(Y,\boldsymbol{\eta})|Y=y_2\right]}{\partial y} \bar{x}^*(y_1) - \frac{\partial \mathbb{E}\left[\pi^d(Y,\boldsymbol{\eta})|Y=y_1\right]}{\partial y} \bar{x}^*(y_2)}{2(x_1 - x_2) y_1 y_2}.$$

The next corollary provides the implication of the independence assumption, i.e., Assumption 3, on Theorem 3.

Corollary 2. *Let Assumptions 1–4 and Data Scheme 2 hold. The partial derivatives of the observed production function with respect to* $\boldsymbol{x}_k, k = 1, \ldots, d_x$ *evaluated at the observed equilibrium matches, i.e.* $\zeta_{\boldsymbol{x}_k}\left(\boldsymbol{x}, \tilde{T}(\boldsymbol{x})^u\right)$ *for* $\boldsymbol{x} \in \mathcal{X}$*, is identified.*

By Theorem 1, Assumptions 3 and 4 imply that the observed type of the downstream partner of each $(\boldsymbol{x}, \boldsymbol{\varepsilon})$-firm is only a function of the observed characteristics $\boldsymbol{x}$. Therefore, the expectation on the right-hand side of (13) can be dropped as the observed equilibrium matching would not include a random term. Namely,

$$\begin{aligned}
\frac{\partial \mathbb{E}\left[\pi^u(\boldsymbol{X}, \boldsymbol{\varepsilon}) | \boldsymbol{X} = \boldsymbol{x}\right]}{\partial x_k} &= \mathbb{E}_{F^u_{\varepsilon|x}}\left[\zeta_{x_k}(\boldsymbol{X}, T^u(\boldsymbol{X}, \boldsymbol{\varepsilon})) | \boldsymbol{X} = \boldsymbol{x}\right] \\
&= \mathbb{E}_{F^u_{\varepsilon|x}}\left[\zeta_{x_k}\left(\boldsymbol{X}, \tilde{T}^u(\boldsymbol{X})\right) | \boldsymbol{X} = \boldsymbol{x}\right] \\
&= \zeta_{x_k}\left(\boldsymbol{x}, \tilde{T}^u(\boldsymbol{x})\right).
\end{aligned}$$

Consequently, the identified term in Corollary 2 is the derivative of the model primitive itself, not a local average of it as in the more general result of Theorem 3.

3.3 Identification of the Nonseparable Model with Profit data

Up to now, we have explored how model assumptions such as independence, separability, and monotonicity in unobserved scalars together with equilibrium matching and profit data enable us to identify the model primitives. In this section, we describe which model primitives can be identified without imposing such assumptions. We consider two assumption scenarios: (i) a non-additively separable production function, non-scalar unobserved characteristics of the firms, and no monotonicity in the unobserved characteristics, whereas maintaining the independence assumption, and (ii) relaxing the independence assumption in the previous case.

The next theorem is an application of Hoderlein and Mammen (2007) and Sasaki (2015) who study the identification of structural partial effects in nonseparable models with multiple unobservables.

Theorem 4. *Let Assumptions 1–3 hold. Let $q_\alpha^u(\boldsymbol{x})$ denote the α-quantile of the upstream firm's profits conditional on $\boldsymbol{X} = \boldsymbol{x}$. That is, for $\alpha \in [0,1]$,*

$$\Pr\left(\pi^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right) \leq q_\alpha^u\left(\boldsymbol{x}\right) | \boldsymbol{X} = \boldsymbol{x}\right) = \alpha.$$

The weighted averages of the production function partial derivatives with respect to the observed characteristics evaluated at the equilibrium matches and conditional on the observed characteristics $\boldsymbol{X} = \boldsymbol{x}$ and the α-quantile of the equilibrium profits, namely,

$$\mathbb{E}_\nu\left[\Phi_{x_k}\left(\boldsymbol{X}, \boldsymbol{\varepsilon}, T^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right), S^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right)\right) | \boldsymbol{X} = \boldsymbol{x}, q_\alpha^u\left(x\right)\right],\ k = 1, 2, \ldots, d_x,$$

where the measure $\nu\left(\boldsymbol{x}, \cdot\right)$ is proportional to $f_\varepsilon^u\left(\cdot\right) / \left\|\nabla_{\boldsymbol{\varepsilon}} \pi^u\left(\boldsymbol{x}, \cdot\right)\right\|$, are identified.

The proof is as follows. Recall (5)

$$\pi^u\left(\boldsymbol{x}, \varepsilon\right) = \max_{(\boldsymbol{y}, \eta) \in \mathcal{Y}} \left\{\Phi\left(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta\right) - \pi^d\left(\boldsymbol{y}, \eta\right)\right\} \equiv m\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right),$$

for some nonseparable function $m(\cdot, \cdot)$. Under the independence assumption[3] and Assumptions 1–2, Theorem 2.1 in Hoderlein and Mammen (2007) implies that for each $\alpha \in [0,1]$,

$$\mathbb{E}_\nu\left[\frac{\partial m\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right)}{\partial x_k} | \boldsymbol{X} = \boldsymbol{x}, \pi^u = q_\alpha^u\left(\boldsymbol{x}\right)\right] = \frac{\partial q_\alpha^u\left(\boldsymbol{x}\right)}{\partial x_k},$$

where $q_\alpha^u(\boldsymbol{x})$ denotes the α_{π^u}-th conditional quantile of the upstream profit conditional on $\boldsymbol{X} = \boldsymbol{x}$ and ν is the measure introduced in Theorem 4 due to Sasaki (2015). The right-hand side is the quantile partial derivative of the upstream profits conditional on the observed characteristics which is identified under Data Scheme 2 and the independence assumption. On the other hand, an application of the envelope theorem to the upstream profit function implies that for each observed characteristic x_k, the partial derivative of the upstream profit with respect to x_k evaluated at $(\boldsymbol{x}, \boldsymbol{\varepsilon})$, can be written as the derivative of the production function with respect to

[3] We can relax this assumption to conditional independence of x_k and $\boldsymbol{\varepsilon}$ given $\boldsymbol{x}_{-k}$.

x_k-argument evaluated at the equilibrium match of $(\boldsymbol{x}, \boldsymbol{\varepsilon})$, i.e.

$$\frac{\partial \pi^u(\boldsymbol{x}, \boldsymbol{\varepsilon})}{\partial x_k} = \Phi_{x_k}\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, T^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right), S^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right)\right),$$

which completes the proof.

Theorem 4 and its proof describe how the partial derivates of the quantiles of the upstream/downstream equilibrium profits conditional on the observed characteristics can be linked to a weighted average of structural partial effects of the production function within the subpopulation of upstream/downstream firms that have the same observed characteristics and equilibrium profit equal to the α-quantile of the upstream/downstream profits conditional on their observed characteristic. Though, the subpopulation of firms included in the average is identified in the data, we cannot identify the weights. Intuitively, larger weights are assigned to those points at which the unobservables have a smaller marginal effect on the structural function. The original result in Hoderlein and Mammen (2007) suggests that the identified feature is a simple average of the partial effects with respect to the underlying measure of the unobserved characteristics. However, Sasaki (2015) revisits and updates Hoderlein and Mammen (2007)'s result by pointing out that the local slope parameter of the quantile regression identifies a weighted average of the underlying structural partial effects, as described in Theorem 4.

Next, we introduce an assumption ensuring that the weights in Theorem 4 are proportional to the measure induced by the underlying distribution of the unobserved characteristics.

Assumption 6. *The gradient norm of the production function with respect to $\boldsymbol{\varepsilon}$, i.e., $\|\nabla_{\boldsymbol{\varepsilon}} \Phi\left(\boldsymbol{x}, \boldsymbol{\varepsilon}, \boldsymbol{y}, \boldsymbol{\eta}\right)\|$), is only a function of $\boldsymbol{x}$.*

An implication of Assumption 6 is that conditional on $\boldsymbol{X} = \boldsymbol{x}$, the denominator of the weight term in Theorem 4 is constant on the set $\{\boldsymbol{\varepsilon} \in \mathcal{E} | \pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) = q^u_\alpha\left(\boldsymbol{x}\right)\}$. The following corollary summarizes this result.

Corollary 3. *Let Assumptions 1–3 and 6 hold. The conditional average partial effects of the structural function*

$$\mathbb{E}\left[\Phi_{x_k}\left(\boldsymbol{X}, \boldsymbol{\varepsilon}, T^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right), S^u\left(\boldsymbol{X}, \boldsymbol{\varepsilon}\right)\right) | \boldsymbol{X} = \boldsymbol{x}, q^u_\alpha\left(x\right)\right], \; k = 1, 2, \ldots, d_x,$$

where the average is over all $(\boldsymbol{x}, \boldsymbol{\varepsilon})$-firms where $\boldsymbol{\varepsilon}$ is such that $\pi^u\left(\boldsymbol{x}, \boldsymbol{\varepsilon}\right) = q^u_\alpha\left(\boldsymbol{x}\right)$.

4 Identification under Price Data Availability

In this section, we explore the identification of the model primitives with data on match transfers instead of firms' profits. Before stating any identification result, we discuss the complexity of using transfer data. We recall that profits are the optimal values of the firm's objective function when choosing a partner. Unlike profits, match transfers are not directly maximized or minimized by firms. For instance, a CEO who has nonpecuniary preferences may choose to manage a firm that is not paying the highest salary but rather the firm that maximizes overall utility, e.g. firm's prestige.

We can express the upstream price function for realized matches in terms of the equilibrium profits and the upstream valuation function as

$$p\left(x,\varepsilon\right)=\pi^{u}\left(x,\varepsilon\right)-\Phi^{u}\left(x,\varepsilon,T^{u}\left(x,\varepsilon\right),S^{u}\left(x,\varepsilon\right)\right).$$

The partial derivative of this price function with respect to the unobserved scalar type is given by

$$\begin{aligned}\frac{\partial p\left(x,\varepsilon\right)}{\partial\varepsilon}=&\pi_{\varepsilon}^{u}\left(x,\varepsilon\right)-\frac{\partial\Phi^{u}\left(x,\varepsilon,T^{u}\left(x,\varepsilon\right),S^{u}\left(x,\varepsilon\right)\right)}{\partial\varepsilon}\\=&\pi_{\varepsilon}^{u}\left(\boldsymbol{x},\varepsilon\right)-\Phi_{\varepsilon}^{u}\left(x,\varepsilon,T^{u}\left(x,\varepsilon\right),S^{u}\left(x,\varepsilon\right)\right)\\&-T_{\varepsilon}^{u}\left(x,\varepsilon\right)\Phi_{y}^{u}\left(x,\varepsilon,T^{u}\left(x,\varepsilon\right),S^{u}\left(x,\varepsilon\right)\right)\\&-S_{\varepsilon}^{u}\left(x,\varepsilon\right)\Phi_{\eta}^{u}\left(x,\varepsilon,T^{u}\left(x,\varepsilon\right),S^{u}\left(x,\varepsilon\right)\right).\end{aligned}\tag{14}$$

In view of (14), assuming the monotonicity of equilibrium transfer would involve assumptions not only on model primitives, but also on model outcomes. The second equality in (14) includes the direct effects of the change in ε on profits and the valuation function, and the indirect effects of the valuation function through the changes in the equilibrium matching partner. Unlike profit function, there is no condition on model primitives that implies monotonicity of the equilibrium matching function. This means that data on prices for realized matches cannot pin down the quantiles of the unobservables.

Before discussing identification of the model primitives with price data, we briefly mention a corollary of Theorem 2 that illustrates how price data in addition to firms' profits identify the valuation functions.

Corollary 4. *Let Assumptions 1,2,3, and 5, and Data Scheme 1 of Theorem 2 hold. We assume prices for all formed matches are observed. Then, the match production* *and valuation functions evaluated at the quantiles of the unobserved distributions are identified on the equilibrium path.*

Corollary 4 uses the fact that profits can be written as the sum of valuation functions and equilibrium transfers. Specifically,

$$\bar{\Phi}^u\left(\boldsymbol{x}, \alpha_\varepsilon, \bar{T}^u\left(\boldsymbol{x}, \alpha_\varepsilon\right), \bar{S}^u\left(\boldsymbol{x}, \alpha_\varepsilon\right)\right) = \bar{\pi}^u\left(\boldsymbol{x}, \alpha_\varepsilon\right) - \bar{p}^u\left(\boldsymbol{x}, \alpha_\varepsilon\right),$$
$$\bar{\Phi}^d\left(\bar{T}^d\left(\boldsymbol{y}, \alpha_\eta\right), \bar{S}^d\left(\boldsymbol{x}, \alpha_\eta\right), \boldsymbol{y}, \alpha_\eta\right) = \bar{\pi}^d\left(\boldsymbol{y}, \alpha_\eta\right) + \bar{p}^d\left(\boldsymbol{x}, \alpha_\eta\right).$$

Since α_ε and α_η are identified under profit monotonicity, the right-hand sides in the previous equations are observed.

As noted in Section 1, the equilibrium matching, profits, and prices do not have a closed form solution in general. However, the well-known result by Dowson and Landau (1982) shows that when the match production is quadratic in the characteristics and the characteristics on each side of the market are jointly normally distributed, the equilibrium admits a closed-form solution. Lindenlaub (2017) uses this result to study worker-job complementarities in an equilibrium matching model. Bojilov and Galichon (2016) extend this result by deriving the closed form while adding logit unobserved heterogeneity to the model similar to Choo and Siow (2006). We use the same model specification and the closed-form expression for equilibrium objects to explore the model identification with price/transfer data.

We first describe the parametric model and then state the identification result.

The Quadratic production and Normal Characteristics Model

Let $\tilde{\boldsymbol{X}} = (x_1, x_2, \varepsilon) \sim N\left(\boldsymbol{0}, \Sigma_{\tilde{X}}\right)$, $\tilde{\boldsymbol{Y}} = (y_1, y_2, \eta) \sim N\left(\boldsymbol{0}, \Sigma_{\tilde{Y}}\right)$, and $\sigma_\varepsilon^2 \equiv Var(\varepsilon) = \sigma_\eta^2 \equiv Var\left(\eta\right) = 1$. Let Assumption 3 hold, i.e., the unobserved characteristics are independently distributed from the observed characteristics on each side of the market. Furthermore, let the match production function as

$$\Phi\left(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta\right) = \beta_{11}x_1y_1 + \beta_{12}x_1y_2 + \beta_{21}x_2y_1 + \beta_{22}^u x_2y_2 + \beta_{13}^u x_1\eta + \beta_{33}^d \varepsilon\eta.$$

The corresponding upstream and downstream valuation functions are given by

$$\Phi^u\left(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta\right) = \beta_{11}^u x_1y_1 + \beta_{12}^u x_1y_2 + \beta_{21}^u x_2y_1 + \beta_{22}^u x_2y_2 + \beta_{13}^u x_1\eta,$$
$$\Phi^d\left(\boldsymbol{x}, \varepsilon, \boldsymbol{y}, \eta\right) = \beta_{11}^d x_1y_1 + \beta_{12}^d x_1y_2 + \beta_{21}^d x_2y_1 + \beta_{33}^d \varepsilon\eta.$$

Note that the production function above is neither additively separable in the observed and unobserved types as in Assumption 4, nor the unobserved terms are separable in the sense of Choo and Siow (2006) due to the presence of $\beta^d_{33}\varepsilon\eta$.

The equilibrium matching function is then a linear function in the characteristics, and can be characterized by the reduced-form

$$\begin{aligned}\begin{pmatrix} y_1 \\ y_2 \\ \eta \end{pmatrix} &= T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix} \\ &= \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix},\end{aligned}$$

where T is the matching matrix given by $T \equiv \Sigma_Y^{1/2}\left(\Sigma_Y^{1/2}A'\Sigma_X A\Sigma_Y^{1/2}\right)^{-1/2}\Sigma_Y^{1/2}A'$, where

$$\Sigma_Y = \begin{pmatrix} \sigma_{y_1} & 0 & 0 \\ 0 & \sigma_{y_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \Sigma_X = \begin{pmatrix} \sigma_{x_1} & 0 & 0 \\ 0 & \sigma_{x_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A = \begin{pmatrix} \beta^d_{11} + \beta^u_{11} & \beta^d_{12} + \beta^u_{12} & \beta^u_{13} \\ \beta^d_{21} + \beta^u_{21} & \beta^u_{22} & 0 \\ \beta^u_{31} & 0 & \beta^d_{33} \end{pmatrix}.$$

The equilibrium transfer of the (x_1, x_2, ε)-firm is given by

$$\begin{aligned}p\left(x_1, x_2, \epsilon\right) &= \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}' B \cdot T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}' \begin{pmatrix} \beta^d_{11} - \beta^u_{11} & \beta^d_{12} - \beta^u_{12} & -\beta^u_{13} \\ \beta^d_{21} - \beta^u_{21} & -\beta^u_{22} & 0 \\ 0 & 0 & \beta^d_{33} \end{pmatrix} \cdot \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}.\end{aligned}$$

Proposition 1. *Let the match production function and the characteristics distribution be given by the quadratic and normal model described above. Let Data Scheme 3 hold. The vector of parameters of the firms' valuation functions* $\boldsymbol{\theta} \equiv \left(\beta^u_{11}, \beta^d_{11}, \beta^u_{12}, \beta^d_{12}, \beta^u_{21}, \beta^d_{21}, \beta^u_{22}, \beta^u_{13}, \beta^d_{33}\right)$ *is identified.*

The proof is provided in the appendix. Although the specification of the model in 1 has been used in empirical research, e.g. Lindenlaub (2017), the model restrictions have implications that can easily be refuted by data. For instance, the assumption that the observed characteristics on the upstream or downstream firms are jointly normally distributed may not be satisfied. Moreover, the model implies that the equilibrium matching, i.e. the outcome of the model, is also normally distributed, i.e. the observed (x_1, x_2, y_1, y_2) are also jointly normally distributed.

Another issue with the normal distribution of the characteristics when the production function is quadratic is the simultaneous negative realization of the upstream and downstream firms' characteristics. In this case, the product of two negative numbers would be the same as the product of two positive values.

Despite the shortcomings of the jointly normal model, it provides evidence that price data are informative about the parameters of the firms' valuation functions. We can expect when the characteristics are not jointly normally distributed, model identification is still maintained. Nevertheless, any deviation from the production function's specification and/or the normality distribution of the characteristics would make the closed-form solutions that are used in Proposition 1 invalid. Currently, we do not have a formal proof of identification that does not rely on the closed-form solutions. Instead, we use Monte Carlo simulations to verify that identification is not violated.

The model we use for the empirical application relaxes the joint normality of characteristics. Specifically, the only restriction on the observed characteristics is their distribution to be continuous as in Assumption 3. We consider a parametric distribution for the unobservables, e.g. log-normal distribution. However, similar to Proposition 1, we normalize the scale of the unobserved characteristics.

5 Estimation

In this section, we propose estimators for the identified model features under the data schemes described in Sections 3 and 4. We first consider a nonparametric estimator to use with profit data. Next, we suggest a simulation-based likelihood estimation procedure to estimate the parametric model with price data.

5.1 Profit Data Estimator

Recall the identification result of Theorem 2. Our goal is to estimate the production function $\bar{\Phi}(\boldsymbol{x}, \alpha_\varepsilon, \boldsymbol{y}, \alpha_\eta)$ at every formed match in the market. Let the sample be given by $\left\{(\boldsymbol{x}_i, \boldsymbol{y}_i^*), \pi_i^u, \pi_i^d\right\}_{i=1}^N$ as in Data Scheme 1. For each upstream firm i in the sample, we need to recover $\alpha_{\varepsilon i}$ as the distribution of upstream profits evaluated at π_i^u and conditional on $\boldsymbol{X} = \boldsymbol{x}$. That is for each observation i, let $\hat{\alpha}_{\varepsilon i} = \hat{F}^u_{\pi^u|x}(\pi_i^u|\boldsymbol{X} = \boldsymbol{x}_i)$, where $\hat{F}^u_{\pi^u|x}(\cdot|\boldsymbol{X} = \boldsymbol{x}_i)$ is a nonparametric estimate of the upstream profit's conditional distribution evaluated at the observed upstream profit of observation i and conditional on the vector of observed characteristics of the upstream firm i.

Similarly, we can recover the quantiles $\alpha^*_{\eta i}$ as the distribution of downstream profits evaluated at the vector of observed characteristics of the downstream matching partner of observation i, i.e. $\hat{\alpha}^*_{\eta i} = \hat{F}^u_{\pi^d|y}\left(\pi^d_i|\boldsymbol{Y} = \boldsymbol{y}^*_i\right)$. Note that

$$F^u_{\pi^u|x}\left(\pi^u|\boldsymbol{X} = \boldsymbol{x}_i\right) = \mathbb{E}\left[\mathbb{1}\left(\pi^u \leq \pi^u_i\right)|\boldsymbol{X} = \boldsymbol{x_i}\right],$$

and

$$F^d_{\pi^d|y}\left(\pi^d|\boldsymbol{Y} = \boldsymbol{y}^*_i\right) = \mathbb{E}\left[\mathbb{1}\left(\pi^d \leq \pi^d_i\right)|\boldsymbol{Y} = \boldsymbol{y^*_i}\right].$$

Thus, the conditional distributions can be estimated by a nonparametric regression of $\mathbb{1}\left(\pi^u \leq \pi^u_i\right)$ and $\mathbb{1}\left(\pi^d \leq \pi^d_i\right)$ on $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively. One estimator would be a simple Nadarya-Watson estimator[4], i.e.

$$\hat{\alpha}_{\varepsilon i} = \hat{F}_{\pi^u|x}\left(\pi^u_i|\boldsymbol{X} = \boldsymbol{x}_i\right) = \frac{\sum_{j=1}^N K_h\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)\mathbb{1}\left(\pi^u_j \leq \pi^u_i\right)}{\sum_{j=1}^n K_h\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)},$$

and similarly for the downstream matching partner of observation i

$$\hat{\alpha}^*_{\eta i} = \hat{F}_{\pi^d|y}\left(\pi^d_i|\boldsymbol{Y} = \boldsymbol{y}^*_i\right) = \frac{\sum_{j=1}^N K_h\left(\boldsymbol{y}^*_i, \boldsymbol{y}^*_j\right)\mathbb{1}\left(\pi^d_j \leq \pi^u_i\right)}{\sum_{j=1}^n K_h\left(\boldsymbol{y}^*_i, \boldsymbol{y}^*_j\right)},$$

where $K_h(\cdot,\cdot)$ is a multivariate kernel. For instance, one can use the normal kernel

$$K_h\left(\boldsymbol{x}_j, \boldsymbol{x}_i\right) = \prod_{d=1}^{d_x}\phi\left(\frac{x_{di} - x_{dj}}{h}\right),$$

where $\phi(\cdot)$ is the standard normal density, and h is a bandwidth. The bandwidth can be chosen by a data driven cross-validation method. For instance, we can minimize the residuals from estimating the empirical CDF and the unconditional CDF estimated as above except that at each observation i, we only use the other $N-1$ sample points to estimate the conditional CDF. To this end, we define the leave-one-out estimator of the conditional CDF as

$$\hat{F}_{\pi^u|x,-i}\left(\pi^u|\boldsymbol{X} = \boldsymbol{x}\right) = \frac{\sum_{j\neq i} K_h\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)\mathbb{1}\left(\pi^u_j \leq \pi^u_i\right)}{\sum_{j\neq i} K_h\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)}.$$

[4]see Li and Racine (2007, p. 182) for details on conditional CDF estimation.

The cross-validation criterion for a fixed profit level π^u is

$$CV\left(\pi^u,h\right)=\frac{1}{n}\sum_{i=1}^{N}\Big(1\left(\pi_i^u\leq\pi^u\right)-\hat{F}_{\pi^u|x,-i}\left(\pi^u|\boldsymbol{X}=\boldsymbol{x}\right)\Big)^2.$$

The optimal bandwidth minimizes

$$CV\left(h\right)=\int CV\left(\pi^u,h_x\right)d\pi^u.$$

We approximate this by a grid over the values of profits, by randomly selecting N_π profit observations. This gives

$$CV\left(h\right)\approx\sum_{i=1}^{N_\pi}CV\left(\pi_i^u,h\right).$$

And , the optimal bandwidth is defined as

$$h^*=\arg\min_h\left\{\sum_{i=1}^{N_\pi}CV\left(\pi_i^u,h\right)\right\}.$$

Example 3. Let $\Phi\left(x,\varepsilon,y,\eta\right)=\beta_{xy}xy+\beta_{x\eta}x\eta+\beta_{y\varepsilon}y\varepsilon+\varepsilon\eta$. Assume the production function is monotone in the unobserved scalars ε and η, i.e.

$$\beta_{y\varepsilon}y+\eta>0,\quad\beta_{x\eta}x+\varepsilon>0.$$

Further, assume that ε and η are independent from x and y, and

$$\varepsilon\sim\text{LogNormal}\left(\mu_\varepsilon,\sigma_\varepsilon\right),\quad\eta\sim\text{LogNormal}\left(\mu_\eta,\sigma_\eta\right).$$

The medians of log-normally distributed ε and η are normalized to one, i.e. $\mu_\varepsilon=\mu_\eta=0$. The equilibrium matching, upstream and downstream profits are observed. Observation i is the tuple $\left(x_i,y_i^*,\pi_i^u,\pi_i^d\right)$. We consider a parameterization of the model and estimate the model parameters. For each observation, $\hat{\alpha}_{\varepsilon i}$ and $\hat{\alpha}_{\eta i}^*$ are estimated according to the procedure explained above. Next, for each choice of the parameters σ_ε and σ_η, we invert $\hat{\alpha}_{\varepsilon i}$ and $\hat{\alpha}_{\eta i}^*$ into the quantiles of the parameterized distributions

$$\hat{\varepsilon}_i=q\left(\hat{\alpha}_{\varepsilon i}|\sigma_\varepsilon\right),\quad\hat{\eta}_i=q\left(\hat{\alpha}_{\eta i}^*|\sigma_\eta\right),$$

		Firms = 500		Firms = 1000		Firms =1500	
	truth	**Bias**	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
β_{xy}	-3.0	0.04	0.48	-0.06	0.29	-0.04	0.26
$\beta_{x\eta}$	0.7	0.02	0.11	0.02	0.09	0.02	0.07
$\beta_{y\varepsilon}$	3.0	-0.06	0.41	0.02	0.25	0.00	0.22
σ_{ε}	0.2	0.02	0.06	0.01	0.05	0.01	0.05
σ_{η}	0.7	0.05	0.08	0.04	0.06	0.04	0.05

Table 1: Estimated Bias and RMSE of the estimator of the model parameters using the equilibrium matching and profit data. We ran a Monte-Carlo experiment with 100 artificial datasets generated from the model in (8) and the parameterization in the table to calculate the average bias and RMSE for three different sample sizes of $N = 500, 1000, 1500$ firms.

where q is the quantile function for the log-normal distribution with $\mu = 0$. Finally, the estimator for $\boldsymbol{\theta} = (\beta_{xy}, \beta_{x\eta}, \beta_{y\varepsilon}, \sigma_{\varepsilon}, \sigma_{\eta})$ minimizes

$$\sum_{i=1}^{n} \left[\left(\pi_i^u + \pi_i^d \right) - \left(\beta_{xy} x_i y_i + \beta_{x\eta} x_i \hat{\eta}_i + \beta_{y\varepsilon} y_i \hat{\varepsilon}_i + \hat{\varepsilon}_i \hat{\eta}_i \right) \right]^2 .$$

Table 5.1 illustrates the performance of the estimator using a Monte Carlo experiment with 100 artificial datasets of equilibrium matching and profits for each market with 500, 1000, and 1500 firms on each side.

5.2 Price Data Estimator

The identification proofs under the availability of profit data suggest appropriate estimation methods, as discussed previously. In the same vein, the parametric identification argument of the quadratic and jointly normal model under the availability of price data in Proposition 1 relies on the closed-form solutions for model outcomes, and the same closed-form solutions allow us to write an expression for the maximum likelihood estimator. In contrast, when the firms' characteristics are not normally distributed, we can neither rely on the identification argument, nor the closed-form expressions for model outcomes to construct an estimator.

We first describe the model specification and discuss the difficulties in estimation. We then propose an estimator for the model parameters. Consider the quadratic production and valuation functions

$$
\begin{aligned}
\Phi\left(\boldsymbol{x},\varepsilon,\boldsymbol{y},\eta\right) &= \beta_{11}x_1y_1+\beta_{12}x_1y_2+\beta_{21}x_2y_1+\beta^u_{22}x_2y_2+\beta^u_{13}x_1\eta+\beta^d_{33}\varepsilon\eta,\\
\Phi^u\left(\boldsymbol{x},\varepsilon,\boldsymbol{y},\eta\right) &= \beta^u_{11}x_1y_1+\beta^u_{12}x_1y_2+\beta^u_{21}x_2y_1+\beta^u_{22}x_2y_2+\beta^u_{13}x_1\eta,\\
\Phi^d\left(\boldsymbol{x},\varepsilon,\boldsymbol{y},\eta\right) &= \beta^d_{11}x_1y_1+\beta^d_{12}x_1y_2+\beta^d_{21}x_2y_1+\beta^d_{33}\varepsilon\eta. \qquad (15)
\end{aligned}
$$

The unobserved characteristics ε and η have a non-negative support and are distributed according to a lognormal distribution with parameters $(\mu = 0, \sigma^2 = 1)$. The marginal distribution of the observed characteristics, $F_x(\cdot)$ and $F_y(\cdot)$, satisfy Assumption 2 but are not assumed to belong to any parametric family. We further assume that the independence assumption, i.e. Assumption 3, holds.

We recall that even though the equilibrium matching for the above model specification is unique, the equilibrium firm profits and transfers are only unique up to an additive constant. In order to pin down the location of profits, we assume that the median downstream profit is equal to $\kappa \in \mathbb{R}$. That is, $\Pr[\pi^d(X,\eta) < \kappa] = 1/2$. Finally, we assume the data satisfy Data Scheme 3, such that the sample can be characterized by $\{x_{1i}, x_{2i}, y^*_{1i}, y^*_{2i}, p_i\}_{i=1}^N$.

We note that the equilibrium matching and prices do not admit a closed-form expression. Consequently, we are not able to write down the log-likelihood function. Furthermore, without knowing the data-generating process, we are not not able to generate data from the model, under different parameter choices, for simulation-based estimation methods.

Nevertheless, the equilibrium of the parametric model above can be numerically approximated. The numerical approximation of optimal transport problems is an active field of research, but we consider a simple and robust approach. The goal is for any choice of parameters $\boldsymbol{\theta} = (\beta^u_{11}, \beta^d_{11}, \beta^u_{12}, \beta^d_{12}, \beta^u_{21}, \beta^d_{21}, \beta^u_{22}, \beta^u_{13}, \beta^d_{33}, \kappa)$, to approximate the data-generating process $g(\cdot,\cdot,\cdot|\boldsymbol{\theta})$ where for every value of $(x_1, x_2, \varepsilon) \in \mathcal{X}$ satisfies

$$
\begin{pmatrix} y^*_1 \\ y^*_2 \\ \eta^* \\ p \end{pmatrix} = g\left(x_1, x_2, \varepsilon|\boldsymbol{\theta}\right). \qquad (16)
$$

The approximation involves solving a discretized version of the continuous problems in (3) and (4), which reduce to the assignment problem in Roth and Sotomayor (1992). In other words, we solve for the equilibrium matching and profits in a finite market where the N_A firms on each side of the market are random draws from the characteristics' distributions $F_{\tilde{x}}(\cdot)$ and $F_{\tilde{y}}(\cdot)$. As N_A increases, the equilibrium matching and profits in the finite market provide a better approximation for the ones in the continuous market. Consider N_A draws of the upstream and downstream firms $\{\boldsymbol{x}_i \equiv (x_{1i}, x_{2i}, \varepsilon_i)\}_{i=1}^{N_A}$ and $\{\boldsymbol{y}_j \equiv (y_{1j}, y_{2j}, \eta_j)\}_{j=1}^{N_A}$, respectively. The equilibrium matching can be characterized by the matrix $\mu^* \in \{0,1\}^{N_A} \times \{0,1\}^{N_A}$ such that firms i and j are matched in equilibrium if and only if $\mu^*_{i,j} = 1$. The finite market equilibrium matching μ^* solves the primal linear programming

$$\max_{\mu \in \{0,1\}^{N_A} \times \{0,1\}^{N_A}} \left\{ \sum_{i=1}^{N_A} \sum_{j=1}^{N_A} \mu_{i,j} \Phi\left(\boldsymbol{x}_i, \varepsilon_i, \boldsymbol{y}_j, \eta_j\right) \right\}$$

subject to

$$\sum_i \mu_{ij} = 1, \text{ for every } j,$$
$$\sum_j \mu_{ij} = 1, \text{ for every } i. \tag{17}$$

The formulation is slightly different from the one in Roth and Sotomayor (1992) in the sense that we do not allow the possibility of staying single, since we do not consider data availability on the unmatched firms. The profits solve the dual of the above linear programming problem given by

$$\min \sum_{i=1}^{N_A} \pi_i^u + \sum_{j=1}^{N_A} \pi_j^d$$

subject to

$$\pi_i^u + \pi_j^d \leq \Phi\left(\boldsymbol{x}_i, \varepsilon_i, \boldsymbol{y}_j, \eta_j\right), \text{ for all } i \text{ and } j. \tag{18}$$

The linear programming problems in (17) and (18) can be solved as a baseline by using linear programming optimizers such as Gurobi. However, there are several algorithms that exploit features of the linear assignment problem to reduce the computational cost and outperform the baseline simplex method for solving the linear program. One of the most popular methods is the one proposed by Jonker and Volgenant (1987) who use the *shortest augmenting paths* based algorithm. We use a variant of this algorithm used by Crouse (2013) to solve for the equilibrium matching and profits in a finite market. Regardless of the algorithm, we denote the solution as $\{x_{1i}, x_{2i}, \eta_i, y^*_{1i}, y^*_{2i}, \eta^*_i, p_i\}_{i=1}^{N_A}$, where i indexes the equilibrium matches and prices.

Next, we return to model estimation. Clearly, the model does not have a tractable likelihood. Also, the moments of the model outcome do not have a closed-form expression. Therefore, we cannot use the traditional likelihood or GMM methods to recover the model parameters. An obvious alternative would be to use a simulation based method such as the Method of Simulated Moments (MSM) proposed by McFadden (1989), where we replace the intractable moments by their simulated counterparts. However, these methods do not lead to an asymptotically efficient estimator, even when the number of simulation draws tends to infinity.

An alternative is to use methods that rely on simulating the likelihood function itself and are asymptotically efficient with an infinite number of simulations. The method we choose was first suggested by Diggle and Gratton (1984), and its asymptotic properties and extension to dynamic models were later studied by Fermanian and Salanié (2004). The nonparametric simulated maximum likelihood method (NPSML) maximizes $L_N(\boldsymbol{\theta}) \equiv (1/N)\sum_{i=1}^{N} \ln l_i(\boldsymbol{\theta})$, where $l_i(\boldsymbol{\theta})$ is the joint density of the outcomes $(y_{1i}^*, y_{2i}^*, p_i)$ given the upstream firm's observed types (x_{1i}, x_{2i}) and parameters vector $\boldsymbol{\theta}$, i.e. $f_{\boldsymbol{y},p}(\boldsymbol{y}_i^*, p_i|\boldsymbol{x}_i;\boldsymbol{\theta})$.

However, as noted earlier, we do not have a closed-form expression for $f_{\boldsymbol{y},p}(\boldsymbol{y}_i^*, p_i| \boldsymbol{x}_i;\boldsymbol{\theta})$. Instead, the NPSML estimator approximates an unknown density by a kernel estimator using a simulated sample generated from the model's reduced form in (16). To this end, for each (x_{1i}, x_{2i}) in the data, we consider S random realizations from the distribution $F_{\varepsilon|x}(\cdot|\boldsymbol{x}_i,\boldsymbol{\theta})$, denoted by $\{\varepsilon_i^s\}_{s=1}^S$. Next, we use (16) or its approximation to compute S model outcomes, namely $\{y_{1i}^{*s}(\boldsymbol{\theta}), y_{2i}^{*s}(\boldsymbol{\theta}), p_i^s(\boldsymbol{\theta})\}_{s=1}^S$. Specifically, we approximate the data-generating process for each simulation draw s by solving the optimization problems in (17) and (18) given the upstream and downstream firms $\{(x_{1i}, x_{2i}, \varepsilon_i^s)\}_{i=1}^N$ and $\{(y_{1j}, y_{2j}, \eta_j^s)\}_{j=1}^N$.

We note that the firms' observed characteristics are not varying across different simulation draws, i.e. they are the same as the observed characteristics in the data. However, the unobserved characteristics are random draws that vary across distributions. Solving for the equilibrium of the finite market simultaneously approximates the data-generating process and also generates the simulated samples required for computing the conditional density. The observed simulated sample is denoted by $\{\{x_{1i}, x_{2i}, y_{1i}^{*s}, y_{2i}^{*s}, p_i^s\}_{i=1}^N\}_{s=1}^S$, where we suppress the argument $\boldsymbol{\theta}$ of the outcomes y_1, y_2, and p. We note that the observed types of the equilibrium downstream partner of the upstream firm i vary across simulations as the unobserved characteristics ε_i^s vary across simulations since they vary with the parameter values $\boldsymbol{\theta}$.

Next, we nonparametrically estimate the conditional density $f_{\boldsymbol{y},p}\left(\boldsymbol{y}_i^*, p_i|\boldsymbol{x}_i;\boldsymbol{\theta}\right)$ as

$$l^S\left(\boldsymbol{y}_i^*, p_i|\boldsymbol{x}_i;\boldsymbol{\theta}\right) \equiv l_i^S\left(\boldsymbol{\theta}\right) \equiv \frac{1}{Sh_{y_1}h_{y_2}h_p}\sum_{s=1}^{S}\phi\left(\frac{y_{1i}^*-y_{1i}^{*s}}{h_{y_1}}\right)\phi\left(\frac{y_{2i}^*-y_{2i}^{*s}}{h_{y_2}}\right)\phi\left(\frac{p_i-p_i^s}{h_p}\right),$$

where $\phi(\cdot)$ is the standard normal density and h_{y_1}, h_{y_2}, h_p are bandwidths for the downstream observed characteristics and price, respectively. We discuss the choice of bandwidths after introducing the estimator. The NPSML estimator $\tilde{\boldsymbol{\theta}}$ is the global maximizer of

$$\begin{aligned}\tilde{L}_N^S\left(\boldsymbol{\theta}\right) &= \frac{1}{N}\sum_{i=1}^{N} l_i^S\left(\boldsymbol{\theta}\right) \\ &= \frac{1}{NSh_{y_1}h_{y_2}h_p}\sum_{i=1}^{N}\sum_{s=1}^{S}\phi\left(\frac{y_{1i}^*-y_{1i}^{*s}}{h_{y_1}}\right)\phi\left(\frac{y_{2i}^*-y_{2i}^{*s}}{h_{y_2}}\right)\phi\left(\frac{p_i-p_i^s}{h_p}\right). \qquad (19)\end{aligned}$$

Equation (19) is a differentiable function of the model parameters $\boldsymbol{\theta}$ if the equilibrium matching and price $y_{1i}^{*s}\left(\boldsymbol{\theta}\right), y_{2i}^{*s}\left(\boldsymbol{\theta}\right), p_i^s\left(\boldsymbol{\theta}\right)$ are differentiable in $\boldsymbol{\theta}$. This is a desired property of the objective function as it allows us to use gradient-based optimization methods to find its maximizers. Even though Assumptions 1–2 ensure differentiability of the outcomes in the models parameters, the finite market approximations are neither differentiable, nor continuous.

To see this, consider a finite market with firms $\{(x_{1i}, x_{2i}, \varepsilon_i)\}_{i=1}^N$and $\{(y_{1j}, y_{2j}, \eta_j)\}_{j=1}^N$. Let $\mu^*(\cdot;\boldsymbol{\theta})$ to be the equilibrium matching in this market under parameters $\boldsymbol{\theta}$ such that $\mu^*\left(i;\boldsymbol{\theta}\right) = j$ if and only if the upstream firm indexed by i is matched to the downstream firm indexed by j in equilibrium. A change in parameters from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}'$ either leave the equilibrium partner of i unchanged if the equilibrium matching is still optimal under $\boldsymbol{\theta}'$, or changes the partner to $j' \neq j$ if the optimal matching is different under $\boldsymbol{\theta}'$. In other words, any change in the parameters may change $y_{1i}^*\left(\boldsymbol{\theta}\right) = y_{1j}$ and $y_{2i}^*\left(\boldsymbol{\theta}\right) = y_{2j}$ to $y_{1i}^*\left(\boldsymbol{\theta}'\right) = y_{1j'}$ and $y_{2i}^*\left(\boldsymbol{\theta}'\right) = y_{2j'}$, where this change in value is not continuous. Since (19) is not differentiable when the equilibrium matching is replaced with a finite market approximation, we need to use derivative-free optimization routines to search for the global maximum.

To evaluate (19) at parameters $\boldsymbol{\theta}$ given an observed sample of size N , we simulate and solve for the equilibrium matching and transfers of S markets of size N. The firms in each market have the same observed characteristics in the data, but unobserved characteristics are random draws from their distributions under $\boldsymbol{\theta}$. In order to pin down the profits in each simulated market, we use shift all the profits so that the median downstream profit in each market is equal to κ, i.e. the equilibrium selection rule which itself is estimated as a parameter of the model.

Choice of Bandwidth

Estimating the conditional density of model outcomes, i.e. $(\boldsymbol{y}^*, p)$, requires choosing a bandwidth for each dimension of the outcome. Using a large bandwidth results in smoothing away some important features of the estimated density, while a small one results in excessive noise that is due to the simulation process. The Silverman (1998) rule-of-thumb is a popular choice and is used in Fermanian and Salanié (2004). However, this would not be an optimal choice if we believe the true density is not a normal distribution. We use the biased cross-validation (BCV) method suggested in Scott and Sain (2005) for choosing the smoothing parameters in multivariate density estimation. The method is based on minimizing an estimate of the asymptotic mean integrated squared error (AMISE) from the data by choosing the bandwidth. The criterion function is given by

$$\begin{aligned} &BCV\left(h_{y1}, h_{y2}, h_p\right) \\ &\quad = \frac{1}{(\sqrt{2\pi})^3 N h_{y1} h_{y2} h_p} + \frac{1}{4N(N-1) h_{y1} h_{y2} h_p} \\ &\quad \times \sum_{i=1}^{n} \sum_{j \neq i} \left[\left(\sum_{k=1}^{3} \Delta_{ijk}^2 \right) - (2 \cdot 3 + 4) \left(\sum_{k=1}^{3} \Delta_{ijk}^2 \right) + \left(3^2 + 2 \cdot 3\right) \right] \\ &\quad \times \prod_{k=1}^{3} \phi\left(\Delta_{ijk}\right), \end{aligned} \tag{20}$$

where $\Delta_{ijk} = {(y^*_{ki} - y^*_{kj})}/{h_{y_k}}$ for $k = 1, 2$ and $\Delta_{ijk} = {(p_i - p_j)}/{h_p}$ for $k = 3$. The optimal choice of bandwidth minimizes (20).

We ran a Monte Carlo experiment to evaluate the estimator's performance using the bandwidths selected by minimizing the biased cross-validation criterion in (20). For each sample size $S = 50, 100, 200$, we generated 100 artificial datasets of observed equilibrium matching and prices using a specific parameterization of the model in (15). We calculated the optimal bandwidths for each artificial dataset and for each outcome dimension by the BCV method. Finally, we estimate the model parameters by maximizing the simulated likelihood function with the simulated sample size equal to the number of observations in the dataset. This choice of simulated sample

		Ups. Firms = 50		Ups. Firms = 100		Ups. Firms =200	
	truth	**Bias**	**RMSE**	**Bias**	**RMSE**	**Bias**	**RMSE**
β^u_{11}	-2.5	0.09	0.71	-0.12	0.63	0.08	0.34
β^u_{12}	1.5	-0.67	0.84	-0.61	0.77	0.7	0.81
β^u_{21}	-1.5	0.35	0.48	0.11	0.37	0.04	0.26
β^u_{22}	-0.5	-0.13	0.43	-0.01	0.22	-0.01	0.19
β^d_{11}	3.5	-1.37	1.87	-1.01	1.26	-0.89	1.21
β^d_{12}	2.5	-1.9	2.01	-1.95	1.99	0.36	1.14
β^d_{21}	1.5	0.80	1.26	0.53	0.72	0.28	0.59
β^u_{13}	3.0	-2.09	-2.52	-1.872	2.22	-1.09	1.31
β^d_{33}	-3.0	2.85	2.92	2.32	2.57	1.15	1.80
κ	3.0	0.965	1.23	0.475	0.83	0.175	0.54

Table 2: Estimated Bias and RMSE of the NPSML estimator of the model parameters using the equilibrium matching and price data. We ran a Monte-Carlo experiment with 100 artificial datasets generated from the model in (15) and the parameterization in the table to calculate the average bias and RMSE for three different sample sizes of $N = 50, 100, 200$ firms. The simulated sample size for each observation equals the corresponding sample size, i.e. $S = N$. The bias and errors decrease with larger sample sizes and simulations.

size implies that the improvement in the estimator's performance is both due to the larger size of the observed sample and the larger size of the simulated sample used to estimate the conditional likelihoods. Moreover, since each simulation is the result of approximating the equilibrium using a finite market of size N, a larger N results in a better approximation of the true data-generating process underlying the continuum matching game

6 Conclusion

In this paper, we set out to explore the empirical significance of the data on two model outcomes, namely profit and price data, in addition to matching data in two-sided matching models where the market is assumed to consist of firms with continuously distributed measured and unmeasured characteristics. To this end, we provide proof for the nonparametric identification of aspects of the structural match production function. Further, we show that price data identifies parameters of match valuation functions of the two sides of the market. On the other hand, although we demonstrate that price data enables us to recover the model primitives in a parametric setting, a more robust nonparametric argument is more complicated than the profit data case.

For data schemes where profit data is available to the researcher, we propose non-parametric estimators directly inspired by the identification argument. In contrast, for price data, we suggest using a simulation-based likelihood estimation method that does not require the model to have a tractable conditional likelihood. The estimation method is straightforward to implement but can be computationally intensive. Also, we need to address computational challenges resulting from the non-differentiability of the simulated likelihood function. Nevertheless, the simulated likelihood estimator is statistically efficient with infinite simulations, which is an advantage compared to other alternatives, such as the method of simulated moments. In both cases, we perform Monte Carlo studies to demonstrate how the performance of these estimators improves with increasing sample size.

While our parametric result provides valuable insight into the identification value of price data, further research is needed to better understand the scope of the non-parametric identification of the model under this data scheme. The computational burden of the simulated likelihood methods makes it a valuable effort to utilize faster algorithms to approximate the solution of the continuum matching game, which is an active area of research. Ideally, a smooth approximation method allows the use of more reliable derivative-based optimization routines, which should significantly decrease the number of evaluations required to find the global maximizers of the simulated likelihood function. Finally, we are using the parametric model presented in this paper to estimate how the general ability of the CEOs, the scope of the firm's operations, and its size affect the match valuation of CEOs and firms using a dataset of large public firms and CEO compensations.

References

Bajari, Patrick and C. Lanier Benkard (2005). "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach". In: *Journal of Political Economy* 113.6, pp. 1239–1276.

Becker, Gary S. (1973). "A Theory of Marriage: Part I". In: *Journal of Political Economy* 81.4, pp. 813–846.

Bojilov, Raicho and Alfred Galichon (2016). "Matching in Closed-Form: Equilibrium, Identification, and Comparative Statics". In: *Economic Theory* 61.4, pp. 587–609.

Carlier, Guillaume (2003). "Duality and Existence for a Class of Mass Transportation Problems and Economic Applications". In: *Advances in Mathematical Economics*. Ed. by Shigeo Kusuoka and Toru Maruyama. Springer, pp. 1–21.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss (2017). "Partner Choice, Investment in Children, and the Marital College Premium". In: *American Economic Review* 107.8, pp. 2109–2167.

Chiappori, Pierre-André et al. (2022). *Assortative Matching on Income*. Working Paper 2022-006.

Choo, Eugene and Aloysius Siow (2006). "Who Marries Whom and Why". In: *Journal of Political Economy* 114.1, pp. 175–201.

Crouse, David Frederic (2013). "Advances in Displaying Uncertain Estimates of Multiple Targets". In: *Signal Processing, Sensor Fusion, and Target Recognition XXII*. Signal Processing, Sensor Fusion, and Target Recognition XXII. Vol. 8745. SPIE, pp. 15–45.

Diggle, Peter J. and Richard J. Gratton (1984). "Monte Carlo Methods of Inference for Implicit Statistical Models". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 46.2, pp. 193–212.

Dowson, D. C and B. V Landau (1982). "The Fréchet Distance between Multivariate Normal Distributions". In: *Journal of Multivariate Analysis* 12.3, pp. 450–455.

Dupuy, Arnaud and Alfred Galichon (2014). "Personality Traits and the Marriage Market". In: *Journal of Political Economy* 122.6, pp. 1271–1319.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim (2004). "Identification and Estimation of Hedonic Models". In: *Journal of Political Economy* 112.S1, S60–S109.

Fermanian, Jean-David and Bernard Salanié (2004). "A Nonparametric Simulated Maximum Likelihood Estimation Method". In: *Econometric Theory* 20.04.

Fox, Jeremy T. (2010). “Identification in Matching Games”. In: *Quantitative Economics* 1.2, pp. 203–254.

— (2018). “Estimating Matching Games with Transfers”. In: *Quantitative Economics* 9.1, pp. 1–38.

Fox, Jeremy T., Chenyu Yang, and David H. Hsu (2018). “Unobserved Heterogeneity in Matching Games”. In: *Journal of Political Economy* 126.4, pp. 1339–1373.

Galichon, Alfred (2016). *Optimal Transport Methods in Economics*. Princeton: Princeton University Press. 170 pp.

Galichon, Alfred and Bernard Salanié (2022). “Cupid’s Invisible Hand: Social Surplus and Identification in Matching Models”. In: *The Review of Economic Studies* 89.5, pp. 2600–2629.

Heckman, James J., Rosa L. Matzkin, and Lars Nesheim (2010). “Nonparametric Identification and Estimation of Nonadditive Hedonic Models”. In: *Econometrica* 78.5, pp. 1569–1591.

Hoderlein, Stefan and Enno Mammen (2007). “Identification of Marginal Effects in Nonseparable Models Without Monotonicity”. In: *Econometrica* 75.5, pp. 1513–1518.

Jonker, R. and A. Volgenant (1987). “A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems”. In: *Computing* 38.4, pp. 325–340.

Li, Qi and Jeffrey Scott Racine (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press. 768 pp.

Lindenlaub, Ilse (2017). “Sorting Multidimensional Types: Theory and Application”. In: *The Review of Economic Studies* 84.2, pp. 718–789.

Matzkin, Rosa L. (2003). “Nonparametric Estimation of Nonadditive Random Functions”. In: *Econometrica* 71.5, pp. 1339–1375.

McFadden, Daniel (1989). “A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration”. In: *Econometrica* 57.5, pp. 995–1026.

Roth, Alvin E. and Marilda Sotomayor (1992). “Chapter 16 Two-sided Matching”. In: *Handbook of Game Theory with Economic Applications*. Vol. 1. Elsevier, pp. 485–541.

Sasaki, Yuya (2015). “What Do Quantile Regressions Identify for General Structural Functions?” In: *Econometric Theory* 31.5, pp. 1102–1116.

Scott, David W. and Stephan R. Sain (2005). “Multidimensional Density Estimation”. In: *Handbook of Statistics*. Vol. 24. Elsevier, pp. 229–261.

Silverman, B. W. (1998). *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability 26. Boca Raton: Chapman & Hall/CRC. 175 pp.

Simon, Carl P. and Lawrence E. Blume (1994). *Mathematics for Economists*. 1st ed. W. W. Norton & Company.

A Proofs

Proof of Theorem 1

Let $\mathcal{F}_{\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Y}}}$ denote the set of all joint distributions over $\tilde{\mathcal{X}} \times \tilde{\mathcal{Y}}$ that have marginals $F^u(\cdot,\cdot)$ and $F^d(\cdot,\cdot)$. Under Assumption 4, the social planner's maximization problem in (3) is given by

$$\max_{F_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}} \in \mathcal{F}_{\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}}} \int_{(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}) \in \tilde{\mathcal{X}} \times \tilde{\mathcal{Y}}} \zeta(\boldsymbol{x},\boldsymbol{y}) + \Xi(\boldsymbol{\varepsilon},\boldsymbol{\eta}) \; dF(\tilde{\boldsymbol{x}},\tilde{\boldsymbol{y}}). \tag{21}$$

The maximization problem in (21) is equivalent to a nested optimization problem, namely

$$\max_{F_{\boldsymbol{x},\boldsymbol{y}} \in \mathcal{F}_{\boldsymbol{x},\boldsymbol{y}}} \int \left[\zeta(\boldsymbol{x},\boldsymbol{y}) + \max_{F_{\boldsymbol{\varepsilon},\boldsymbol{\eta}|\boldsymbol{x},\boldsymbol{y}} \in \mathcal{F}_{\boldsymbol{\varepsilon},\boldsymbol{\eta}|\boldsymbol{x},\boldsymbol{y}}} \int \Xi(\boldsymbol{\varepsilon},\boldsymbol{\eta}) \; dF_{\boldsymbol{\varepsilon},\boldsymbol{\eta}|\boldsymbol{x},\boldsymbol{y}} \right] dF_{\boldsymbol{x},\boldsymbol{y}}, \tag{22}$$

where $\mathcal{F}_{\boldsymbol{x},\boldsymbol{y}}$ is the set of all joint distributions over $\mathcal{X} \times \mathcal{Y}$ with marginals $F^u_{\boldsymbol{x}}(\cdot)$ and $F^d_{\boldsymbol{y}}(\cdot)$, and $\mathcal{F}_{\boldsymbol{\varepsilon},\boldsymbol{\eta}|\boldsymbol{x},\boldsymbol{y}}$ is the set of all joint distributions over $\mathcal{E} \times \mathcal{H}$ that have marginals $F^u_{\boldsymbol{\varepsilon}|\boldsymbol{x}}(\cdot)$ and $F^d_{\boldsymbol{\eta}|\boldsymbol{y}}(\cdot)$.

Under Assumption 3, $\mathcal{F}_{\boldsymbol{\varepsilon},\boldsymbol{\eta}|\boldsymbol{x},\boldsymbol{y}}$ is constant over $\mathcal{X} \times \mathcal{Y}$ as the unobserved types are distributed independently from the observed types. In other words, any distributions over the unobserved types that coincides with the conditional marginal distributions induced by $F^u(\cdot,\cdot)$ and $F^d(\cdot,\cdot)$ is also in the set of joint distributions with *unconditional* marginal distributions. Thus, we can rewrite (22) as

$$\max_{F_{\boldsymbol{x},\boldsymbol{y}} \in \mathcal{F}_{\boldsymbol{x},\boldsymbol{y}}} \int [\zeta(\boldsymbol{x},\boldsymbol{y}) +] \; dF_{\boldsymbol{x},\boldsymbol{y}} + \max_{F_{\boldsymbol{\varepsilon},\boldsymbol{\eta}} \in \mathcal{F}_{\boldsymbol{\varepsilon},\boldsymbol{\eta}}} \int \Xi(\boldsymbol{\varepsilon},\boldsymbol{\eta}) \; dF_{\boldsymbol{\varepsilon},\boldsymbol{\eta}}.$$

Therefore, the equilibrium matching consists of solving two independent maximization problems. The first one is maximizing the expected observed production by choosing a distribution over the observed types, and the second one where the unobserved production is maximized by choosing a joint distribution over the unobserved types. The solutions can be characterized by $(\tilde{T}^u(\cdot), \tilde{S}^u(\cdot))$ where $\tilde{T}^u : \mathbb{R}^{d_x} \mapsto \mathbb{R}^{d_y}$ and $\tilde{S}^u : \mathbb{R}^{d_\epsilon} \mapsto \mathbb{R}^{d_\eta}$, respectively.

Proof of Theorem 2

Consider the sample under Data Scheme 2, denoted by $\{\boldsymbol{x}_i, \boldsymbol{y}^*_i, \pi^u_i, \pi^d_i\}_{i=1}^N$. Under the monotonicity assumption, ${\partial \pi^u(\boldsymbol{x},\varepsilon)}/{\partial \varepsilon} > 0$ and ${\partial \pi^d(\boldsymbol{y},\eta)}/{\partial \eta} > 0$. Consequently,

${}^{\partial \bar{\pi}^u(\boldsymbol{x},\alpha_\varepsilon)}\!/_{\partial \alpha_\varepsilon} > 0$ and ${}^{\partial \bar{\pi}^d(\boldsymbol{y},\alpha_\eta)}\!/_{\partial \alpha_\eta} > 0$. Thus, under the independence assumption we can recover $\alpha_{\varepsilon i}$ and $\alpha_{\eta i}$ as the conditional $\alpha_{\varepsilon i}$- and $\alpha_{\eta i}$-quantiles of the upstream and downstream profits, given $\boldsymbol{X} = \boldsymbol{x}_i$ and $\boldsymbol{Y} = \boldsymbol{y}_i^*$, respectively. That is for each observation i, $\alpha_{\varepsilon i} = \Pr(\pi^u(\boldsymbol{X}, \varepsilon) \leq \pi_i^u | \boldsymbol{X} = \boldsymbol{x}_i)$ and $\alpha_{\eta i} = \Pr(\pi^d(\boldsymbol{Y}, \eta) < \pi_i^d | \boldsymbol{Y} = \boldsymbol{y}_i^*)$, which are both identified under the independence assumption. We can now characterize each observation by the match, firm profits, and the quantiles $\alpha_{\varepsilon i}$ and $\alpha_{\eta i}$, namely $(\boldsymbol{x}_i, \boldsymbol{y}_i, \pi_i^u, \pi_i^d, \alpha_{\varepsilon i}, \alpha_{\eta i})$. Finally, (2) implies that the sum of the profits of matched firms equals the production function at the match, ie. $\bar{\Phi}(\boldsymbol{x}, \alpha_\varepsilon, \boldsymbol{y}, \alpha_\eta) = \pi_i^u + \pi_i^d$.

Proof of Proposition 1

The equilibrium matching is a linear function of the characteristics, such that

$$\begin{aligned} y_1 &= t_{11}x_1 + t_{12}x_2 + t_{13}\varepsilon, \\ y_2 &= t_{21}x_1 + t_{22}x_2 + t_{23}\varepsilon, \\ \eta &= t_{31}x_1 + t_{32}x_2 + t_{33}\varepsilon, \end{aligned}$$

where t_{ij} are elements of the reduced-form matching matrix T which is a non-trivial function of the model parameters. Since the equilibrium matching is determined by the match production function and the characteristics distributions, matrix T is a function of the following seven parameters of the production function $\left(\beta_{11}^d + \beta_{11}^u\right), \left(\beta_{12}^d + \beta_{12}^u\right), \left(\beta_{21}^d + \beta_{21}^u\right), \beta_{22}^u, \beta_{13}^u, \beta_{33}^d$, and the distribution parameters which are either identified for the observed ones, or normalized for the unobserved ones. Since $x_1, x_2 \perp \varepsilon$, regressing y_1 on x_1 and x_2 identifies t_{11}, t_{12} as the coefficients, and $|t_{13}|$ as the standard deviation of the residuals. Similarly, regressing y_2 on x_1 and x_2 identifies t_{21}, t_{22} as coefficients, and $|t_{23}|$ as the standard deviation of the residuals.

The equilibrium matching implies that the following equality

$$\begin{aligned} Cov\left(y_1, y_2\right) &= Cov\left(t_{11}x_1 + t_{12}x_2 + t_{13}\epsilon, t_{21}x_1 + t_{22}x_2 + t_{23}\varepsilon\right) \\ &= t_{11}t_{21}Var\left(x_1\right) + t_{11}t_{22}Cov\left(x_1, x_2\right) + t_{11}t_{23}Cov\left(x_1, \varepsilon\right) + \\ &+ t_{12}t_{21}Cov\left(x_1, x_2\right) + t_{12}t_{22}Var\left(x_2\right) + t_{12}t_{23}Cov\left(x_2, \varepsilon\right) + \\ &+ t_{13}t_{21}Cov\left(\varepsilon, x_1\right) + t_{13}t_{22}Cov\left(\varepsilon, x_2\right) + t_{13}t_{23}Var\left(\varepsilon\right) = \\ &= t_{11}t_{21}\sigma_{x_1}^2 + \left(t_{11}t_{22} + t_{12}t_{21}\right)\sigma_{x_1x_2} + t_{12}t_{22}\sigma_{x_2}^2 + \boldsymbol{t_{13}t_{23}} = \sigma_{y_1y_2}. \end{aligned}$$

The sign of $t_{13}t_{23}$ is identified, since

$$t_{13}t_{23} = \sigma_{y_1y_2} - \left(t_{11}t_{21}\sigma^2_{x_1} + \left(t_{11}t_{22} + t_{12}t_{21}\right)\sigma_{x_1x_2} + t_{12}t_{22}\sigma^2_{x_2}\right),$$

where all the terms on the right-hand side are identified previously.

Further, the scale normalization of η implies

$$\begin{aligned} Var\left(\eta\right) &= Var\left(t_{31}x_1 + t_{32}x_2 + t_{33}\varepsilon\right) \\ &= t^2_{31}\sigma^2_{x_1} + t^2_{32}\sigma^2_{x_2} + t^2_{33} = 1, \end{aligned}$$

where t_{31}, t_{32}, t_{33} are unknown parameters.

And, under Assumption 3, $Cov\left(y_j, \eta\right) = 0$. So,

$$\begin{aligned} &Cov\left(t_{11}x_1 + t_{12}x_2 + t_{13}\epsilon, t_{31}x_1 + t_{32}x_2 + t_{33}\varepsilon\right) = \\ &= t_{11}t_{31}Var\left(x_1\right) + t_{11}t_{32}Cov\left(x_1, x_2\right) + t_{11}t_{33}Cov\left(x_1, \varepsilon\right) + \\ &+ t_{12}t_{31}Cov\left(x_1, x_2\right) + t_{12}t_{32}Var\left(x_2\right) + t_{12}t_{33}Cov\left(x_2, \varepsilon\right) + \\ &+ t_{13}t_{31}Cov\left(\epsilon, x_1\right) + t_{13}t_{32}Cov\left(\epsilon, x_2\right) + t_{13}t_{33}Var\left(\varepsilon\right) = \\ &= t_{11}t_{31}\sigma^2_{x_1} + \left(t_{11}t_{32} + t_{12}t_{31}\right)\sigma_{x_1x_2} + t_{12}t_{32}\sigma^2_{x_2} + t_{13}t_{33} = 0. \end{aligned}$$

Now, consider $Cov\left(y_2, \eta\right)$

$$\begin{aligned} &Cov\left(t_{21}x_1 + t_{22}x_2 + t_{23}\epsilon, t_{31}x_1 + t_{32}x_2 + t_{33}\epsilon\right) = \\ &= t_{21}t_{31}Var\left(x_1\right) + t_{21}t_{32}Cov\left(x_1, x_2\right) + t_{21}t_{33}Cov\left(x_2, \epsilon\right) + \\ &+ t_{22}t_{31}Cov\left(x_1, x_2\right) + t_{22}t_{32}Var\left(x_2\right) + t_{22}t_{33}Cov\left(x_2, \epsilon\right) + \\ &+ t_{33}t_{31}Cov\left(\epsilon, x_1\right) + t_{33}t_{32}Cov\left(\epsilon, x_2\right) + t_{33}t_{33}Var\left(\epsilon\right) = \\ &= t_{21}t_{31}\sigma^2_{x_1} + \left(t_{21}t_{32} + t_{22}t_{31}\right)\sigma_{x_1x_2} + t_{22}t_{32}\sigma^2_{x_2} + t_{23}t_{33} = 0. \end{aligned}$$

Collecting the three equations,

$$\begin{cases} \boldsymbol{t^2_{31}}\sigma^2_{x_1} + \boldsymbol{t^2_{32}}\sigma^2_{x_2} + \boldsymbol{t^2_{33}} = 1, \\ t_{11}\boldsymbol{t_{31}}\sigma^2_{x_1} + \left(t_{11}\boldsymbol{t_{32}} + t_{12}\boldsymbol{t_{31}}\right)\sigma_{x_1x_2} + t_{12}\boldsymbol{t_{32}}\sigma^2_{x_2} + t_{13}\boldsymbol{t_{33}} = 0, \\ t_{21}\boldsymbol{t_{31}}\sigma^2_{x_1} + \left(t_{21}\boldsymbol{t_{32}} + t_{22}\boldsymbol{t_{31}}\right)\sigma_{x_1x_2} + t_{22}\boldsymbol{t_{32}}\sigma^2_{x_2} + t_{23}\boldsymbol{t_{33}} = 0, \end{cases}$$

where the bold variable are the unknowns.

Let us assume $t_{33} \geq 0$, then

$$t_{33} = \sqrt{1 - t_{31}^2 \sigma_{x_1}^2 - t_{32}^2 \sigma_{x_2}^2}. \tag{23}$$

Then, we end up with a system of two (non-linear) equations with two unknowns

$$\begin{cases} t_{11}\boldsymbol{t_{31}}\sigma_{x_1}^2 + \left(t_{11}\boldsymbol{t_{32}} + t_{12}\boldsymbol{t_{31}}\right)\sigma_{x_1 x_2} + t_{12}\boldsymbol{t_{32}}\sigma_{x_2}^2 + t_{13}\sqrt{1 - \boldsymbol{t_{31}^2}\sigma_{x_1}^2 - \boldsymbol{t_{32}^2}\sigma_{x_2}^2} = 0, \\ t_{21}\boldsymbol{t_{31}}\sigma_{x_1}^2 + \left(t_{21}\boldsymbol{t_{32}} + t_{22}\boldsymbol{t_{31}}\right)\sigma_{x_1 x_2} + t_{22}\boldsymbol{t_{32}}\sigma_{x_2}^2 + t_{23}\sqrt{1 - \boldsymbol{t_{31}^2}\sigma_{x_1}^2 - \boldsymbol{t_{32}^2}\sigma_{x_2}^2} = 0. \end{cases} \tag{24}$$

We can solve the system of equations above for the two unknowns. Now, given the matching matrix T, we move onto the prices.

Recall the equilibrium price function

$$p\left(x_1, x_2, \varepsilon\right) = \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}' \underbrace{\begin{pmatrix} \beta_{11}^d - \beta_{11}^u & \beta_{12}^d - \beta_{12}^u & -\beta_{13}^u \\ \beta_{21}^d - \beta_{21}^u & -\beta_{22}^u & 0 \\ 0 & 0 & \beta_{33}^d \end{pmatrix}}_{B} \cdot \underbrace{\begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix}}_{T} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}.$$

Every quadratic form $x'Dx$ for some matrix D, can be equivalently represented by a symmetric matrix given by $\frac{D+D'}{2}$. Therefore, we can rewrite the equivalent form of (15) as

$$p\left(x_1, x_2, \epsilon\right) = \frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}' \cdot \frac{B \cdot T + T'B'}{2} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \varepsilon \end{pmatrix}.$$

Let us define the symmetric matrix characterizing the price quadratic form as $K \equiv \frac{B \cdot T + T'B'}{2}$. K is characterized by only 6 values instead of 9 since it is a symmetric matrix, i.e.

$$K = \begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{12} & K_{22} & K_{23} \\ K_{13} & K_{23} & K_{33} \end{pmatrix},$$

whereas

$$
\begin{aligned}
K_{11} =& 2\left(\beta_{11}^{d}-\beta_{11}^{u}\right)t_{11}+2\left(\beta_{12}^{d}-\beta_{12}^{u}\right)t_{21}-2\beta_{13}^{u}t_{31},\\
K_{12} =& \left(\beta_{21}^{d}-\beta_{21}^{u}\right)t_{11}-\beta_{22}^{u}t_{21}+\left(\beta_{11}^{d}-\beta_{11}^{u}\right)t_{12}+\left(\beta_{12}^{d}-\beta_{12}^{u}\right)t_{22}-\beta_{13}^{u}t_{32},\\
K_{13} =& \beta_{33}^{d}t_{31}+\left(\beta_{11}^{d}-\beta_{11}^{u}\right)t_{13}+\left(\beta_{12}^{d}-\beta_{12}^{u}\right)t_{23}-\beta_{13}^{u}t_{33},\\
K_{22} =& 2\left(\beta_{21}^{d}-\beta_{21}^{u}\right)t_{12}-2\beta_{22}^{u}t_{22},\\
K_{23} =& \beta_{33}^{d}t_{32}+\left(\beta_{21}^{d}-\beta_{21}^{u}\right)t_{13}-\beta_{22}^{u}t_{23},\\
K_{33} =& 2\beta_{33}^{d}t_{33}.
\end{aligned}
$$

Since the elements of T are identified, we can uniquely solve the above system of 6 linear equations for

$$
\left(\beta_{11}^{d}-\beta_{11}^{u}\right),\left(\beta_{12}^{d}-\beta_{12}^{u}\right),\left(\beta_{21}^{d}-\beta_{21}^{u}\right),\beta_{22}^{u},\beta_{13}^{u},\beta_{33}^{d}.
$$

This already proves that the structural parameters $\beta_{22}^{u},\beta_{13}^{u},\beta_{33}^{d}$ are identified. However, this does not separately identify the coefficients of valuation functions, but only their differences.

The matrix T can be written as

$$
\begin{aligned}
T &= \overbrace{\Sigma_Y^{1/2}\left(\Sigma_Y^{1/2}A'\Sigma_X A\Sigma_Y^{1/2}\right)^{-1/2}\Sigma_Y^{1/2}}^{H}A'\\
&=\begin{pmatrix} H_{11} & H_{12} & H_{13}\\ H_{12} & H_{22} & H_{23}\\ H_{13} & H_{23} & H_{33}\end{pmatrix}\begin{pmatrix} \beta_{11}^{d}+\beta_{11}^{u} & \beta_{21}^{d}+\beta_{21}^{u} & \beta_{13}^{u}\\ \beta_{12}^{d}+\beta_{12}^{u} & \beta_{22}^{u} & 0\\ 0 & 0 & \beta_{33}^{d}\end{pmatrix}=\begin{pmatrix} t_{11} & t_{12} & t_{13}\\ t_{21} & t_{22} & t_{23}\\ t_{31} & t_{32} & t_{33}\end{pmatrix}. \qquad (25)
\end{aligned}
$$

Equation (25) exploits the structure of matrix T to express it as the product of a symmetric matrix H and the matrix A which is the sum of upstream and downstream coefficients. Since H is a symmetric matrix, it is only characterized by 6 parameters. Recall, that all 9 elements of matrix T are identified from data on matching, and further the three parameters $\beta_{22}^{u},\beta_{13}^{u},\beta_{33}^{d}$ are identified from the argument above using the price equation. Equation (25) defines a system of 9 equations (not all are linear):

$$
\begin{aligned}
H_{11}\left(\beta_{11}^{u}+\beta_{11}^{d}\right)+H_{12}\left(\beta_{12}^{u}+\beta_{12}^{d}\right) &= t_{11} \\
H_{11}\left(\beta_{21}^{u}+\beta_{21}^{d}\right)+H_{12}\beta_{22}^{u} &= t_{12} \\
H_{11}\beta_{13}^{u}+H_{13}\beta_{33}^{d} &= t_{13} \\
H_{12}\left(\beta_{11}^{u}+\beta_{11}^{d}\right)+H_{22}\left(\beta_{12}^{u}+\beta_{12}^{d}\right) &= t_{21} \\
H_{12}\left(\beta_{21}^{u}+\beta_{21}^{d}\right)+H_{22}\beta_{22}^{u} &= t_{22} \\
H_{12}\beta_{13}^{u}+H_{23}\beta_{33}^{d} &= t_{23} \\
H_{13}\left(\beta_{11}^{u}+\beta_{11}^{d}\right)+H_{23}\left(\beta_{12}^{u}+\beta_{12}^{d}\right) &= t_{31} \\
H_{13}\left(\beta_{21}^{u}+\beta_{21}^{d}\right)+H_{23}\beta_{22}^{u} &= t_{32} \\
H_{13}\beta_{13}^{u}+H_{33}\beta_{33}^{d} &= t_{33}
\end{aligned}
$$

There are 9 unknowns in the above equations:

$$H_{11}, H_{12}, H_{13}, H_{22}, H_{23}, H_{33}, \left(\beta_{21}^{u}+\beta_{21}^{d}\right), \left(\beta_{12}^{u}+\beta_{12}^{d}\right), \left(\beta_{12}^{d}+\beta_{12}^{u}\right).$$

We can solve the system of 9 equations for the 9 unknowns listed above. Collecting the two intermediate identification arguments above using the price equation and the matching equation, we have identified $\left(\beta_{21}^{u}+\beta_{21}^{d}\right), \left(\beta_{12}^{u}+\beta_{12}^{d}\right), \left(\beta_{12}^{d}+\beta_{12}^{u}\right), \left(\beta_{11}^{d}-\beta_{11}^{u}\right), \left(\beta_{12}^{d}-\beta_{12}^{u}\right)$ Since we have identified $\left(\beta_{21}^{u}+\beta_{21}^{d}\right), \left(\beta_{21}^{d}-\beta_{21}^{u}\right)$, we can identify $\beta_{21}^{u}, \beta_{21}^{d}$ separately. Similarly, we can identify $\beta_{12}^{u}, \beta_{12}^{d}, \beta_{11}^{u}, \beta_{11}^{d}$ separately.