Estimating Matching Games with Transfers

Jeremy T. Fox*
University of Michigan and NBER

December 2010

Abstract

I explore the estimation of matching games. I use data on the car parts supplied by automotive suppliers to estimate the returns from different portfolios of parts. I answer questions relevant to policy debates about divesting brands from global parent corporations and encouraging foreign producers to assemble cars domestically. I estimate the structural revenue functions of car parts suppliers and automotive assemblers by imposing that the portfolios of car parts represent a pairwise stable equilibrium to a many-to-many, transferable utility matching game. The maximum score estimator does not suffer from a computational curse of dimensionality in the number of firms in a matching market.

*Thanks to SupplierBusiness as well as Thomas Klier for help with the work on automotive supplier specialization. I thank the National Science Foundation, the NET Institute, the Olin Foundation, and the Stigler Center for generous funding. Thanks to helpful comments from colleagues, referees and workshop participants at many universities and conferences. Chenchuan Li, David Santiago, Louis Serranito and Chenyu Yang provided excellent research assistance. Email: jtfox@umich.edu.

1 Introduction

There are many situations in which economists have data on relationships, including marriages between men and women and partnerships between upstream and downstream firms. Economists wish to use the data on the set of realized relationships to estimate the preferences of agents over the characteristics of potential partners. This is a challenging task compared to estimating preferences using more traditional data because we observe only the equilibrium relationships and not each agent's choice set: the identity of the other agents who would be willing to match with a particular agent. This paper models the formation of relationships as a pairwise stable equilibrium to a two-sided, many-to-many matching game with transferable utility. Using this structure, the paper explores the estimation of structural revenue functions, which represent the preferences of upstream firms for downstream firms and of downstream firms for upstream firms. Computational challenges are key in matching and a computationally simple maximum score estimator is introduced to address those problems. The paper uses the maximum score estimator to empirically answer questions related to automotive parts suppliers. I first describe the empirical application and then the methodological contribution.

A car is one of the most complex goods that an individual consumer will purchase. Cars are made up of hundreds of parts and the performance of the supply chain is critical to the performance of automobile assemblers and the entire industry. This paper investigates two related questions that are relevant to policy debates on the automobile industry. The first question relates to the productivity loss to suppliers from breaking up large assemblers of automobiles. Recently, North American-based automobile assemblers have gone through a period of financial distress. As a consequence, North American-based assemblers have divested or closed both domestic brands (General Motor's Saturn) and foreign brands (Ford's Volvo) and have seriously considered the divestment of other brands (GM's large European subsidiary Opel). One loss from divesting a brand is that future product development will no longer be coordinated across as many brands under one parent company. If GM were to divest itself of Opel, which was a serious policy debate in Germany in 2009, then any benefit from coordinated new products across Opel and GM's North American operations would be lost. This is a loss to GM, but also to the suppliers of GM, who will no longer be able to gain as much from specializing in supplying GM. I will estimate the relative benefits to suppliers and to assemblers for different portfolios of car parts.

The second question this paper investigates is the extent to which the presence of foreign and in particular Japanese and Korean (Asian) assemblers in North America improves the North American supplier base. There is a general perception, backed by studies that I cite, that Asian automobile assemblers produce cars of higher quality. Part of producing a car of higher quality is sourcing car parts of higher quality. Therefore, Asian assemblers located in North America might improve North American suppliers' qualities. Understanding the role of foreign entrants on the North American supplier base is important for debates about trade barriers that encourage Asian assemblers to locate plants in North America in order to avoid those barriers. Trade barriers might indirectly benefit North American assemblers by encouraging higher quality North American suppliers to operate in order to supply Asian-owned assembly plants in North America.

I answer both of the above questions using a relatively new type of data: the identities of the companies that supply each car part. I use a dataset listing each car model and each car part on that model, and importantly the supplier of each car part. The intuition behind my approach is that the portfolio of car parts that each supplier manufactures tells us a lot about the factors that make a successful supplier. If each supplier sells car parts to only two assemblers, it may be that suppliers benefit from specialization at the assembly firm level. If North American suppliers to Asian-owned assemblers are also likely to supply parts to North American-owned assemblers, it may be because of a competitive, quality advantage that those suppliers have.

This paper takes the stand that the sorting pattern of upstream firms (suppliers like Bosch and Delphi) to downstream firms (automobile assemblers like General Motors and Toyota) can inform us about the structural revenue functions, key components of total profits, generating the payoffs of particular portfolios of car part matches to suppliers and to assemblers. In turn, the revenue function for suppliers and the revenue function for assemblers help us answer the policy questions about government-induced divestitures and foreign assembler plants in North America.

I will need to introduce an appropriate theoretical framework in order to use data on the identity of car parts suppliers for particular car models in a revealed preference approach to estimate, up to scale, the structural revenue function for a portfolio of car parts. I model the market for car parts as a two-sided matching market, with the two sides being suppliers and assemblers. In this matching market, suppliers are rivals to sells parts to assemblers and assemblers may be rivals to match with the best suppliers. Each firm will form the matches, car part transactions, that maximize its profits at the market-clearing prices. However, those prices are not in my data; they are confidential contractual details not released to researchers. So my revealed preference approach will need more than the individual rationality condition that firms maximize profits given the prices they are paying or being charged. I will take an explicit stand on the equilibrium being played in the matching market for car parts. I will assume that the matches between suppliers and assemblers in the data represent an equilibrium outcome that is pairwise stable, which I will define.

A critical feature of the two policy questions that I will answer is that they involve the structural revenue functions that give the net revenue (implicitly subtracting costs) from the portfolios of car part matches made by suppliers and by assemblers. The loss to a supplier from GM divesting Opel occurs when supplying two car parts to a large parent company generates more revenue than supplying one car part each to two car companies. Thus, this paper works with structural revenue functions that are not the sums of the revenue from individual car part matches. Revenue functions are not additively separable across multiple matches, as they are in some prior work on a different type of matching game (one without money), such as Sørensen (2007). Compared to Fox (2010) and Sørensen (2007), I show how to separately estimate the structural revenue functions of upstream firms and of downstream firms. I can distinguish the payoffs of one side of the market from the payoffs of the other side. As I explain in the text, this is only possible in many-to-many matching; in one-to-one matching my identification strategy would typically only identify the sum of the revenue functions from both sides of the market.

In terms of matching theory, I model the markets for car parts as two-sided, many-to-many matching games with transferable utility. The "two sides" are the suppliers and assemblers. "Many-to-many" means that each assembler has multiple suppliers and each supplier sells to multiple assemblers. "Transferable utility" means each assembler gives money to its suppliers, and both assemblers and suppliers express their utilities in terms of money. Transferable utility is a reasonable assumption for many firms. This framework of matching for modeling upstream and downstream firm relationships can be extended to other industries, for example the matching of manufacturers or distributors of goods to retailers, taking into account shelf-space constraints. Another example is the one-to-many matching of mobile phone carriers to geographic spectrum licenses in an FCC spectrum auction, which I study using the techniques in this paper in Fox and Bajari (2010). Two-sided, many-to-many matching games are a generalization of many other special cases, including the one-sided matching of firms to other firms in mergers, the one-to-many, two-sided matching of workers to firms, and the one-to-one, two-sided matching of men to women in marriage. The methods in this paper can be applied to these other matching markets as well.

Computational issues in matching games are paramount and, in my opinion, have limited the prior use of matching games in empirical work. Matching markets often have hundreds of firms in them, compared to the two to four firms often modeled as potential entrants in applications of Nash entry games in industrial organization. In the car parts data, there are 2627 car parts in one particular car component category. Because of the history of the automotive supplier industry, I treat each component category as a separate matching market. There are thus 2627 opportunities for a car parts supplier to match with an assembler in a single matching market. In Fox and Bajari (2010), we apply a related version of the estimator in this paper to the matching between bidders and items for sale in a FCC spectrum auction. There are 85 winning bidders and 480 items for sale in the auction application. Both the automotive supplier and auction datasets are rich. There is a lot of information on agent characteristics and a lot of unknown parameters that can be learned from the observed sorting of suppliers to assemblers or bidders to items for sale. To take advantage of rich data sets, a researcher must propose an estimator that works around the dimensionality of typical problems.

This paper introduces a computationally simple, maximum score estimator for structural revenue functions (Manski, 1975, 1985; Horowitz, 1992; Matzkin, 1993; Fox, 2007; Jun et al., 2009). The estimator uses inequalities derived from necessary conditions for pairwise stability. There is a tradition of using necessary conditions or inequalities to estimate complex games. See Haile and Tamer (2003) and Bajari, Benkard and Levin (2007) for applications to noncooperative, Nash games. In my estimator, these necessary conditions involve only observable firm characteristics; there is no potentially high-dimensional integral over unobservable characteristics. Evaluating the statistical objective function is computationally simple: checking whether an inequality is satisfied requires only evaluating revenue functions and conducting pairwise comparisons. The objective function is the number of inequalities that are satisfied for any guess of the structural parameters. The estimators are any parameters for the two revenue functions that maximize the number of included inequalities. Because the set of inequalities can be large, I argue that the estimator will be consistent if the researcher samples from the set of possible inequalities. Numerically computing the global maximum of the objective function requires a global optimization routine, although estimation is certainly doable with software built into commercial packages such as MATLAB or Mathematica. Some effort must be spent on running the optimization software multiple times to check the robustness of the optimum. A Monte Carlo study illustrates the important computational advantages of the maximum score estimator by comparing its performance on seemingly trivial matching estimation problems to two parametric, simulation estimators.

The estimator is semiparametric as the structure on unobservables is not modeled up to a finite vector of parameters. Indeed, the maximum score estimator is consistent because of a rank order property that relates the inequalities from pairwise stability to the probabilities of different equilibrium assignments. I introduce two rank order properties corresponding to different asymptotic arguments: collecting data on more independent matching markets or more data on one large matching market. The first rank order property leads to a maximum score estimator like Manski (1975) and the second rank order property leads to a maximum rank correlation estimator like Han (1987).

For one-to-one matching, a sufficient condition for the rank order property for the one large matching market is the set of assumptions underlying the logit based matching model of marriage of Choo and Siow (2006), the only prior paper on estimating transferable utility matching games.[1] Therefore, the model considered in this paper strictly generalizes prior work by allowing for the logit errors in Choo and Siow but not imposing them. For many-to-many matching games where the structural revenue function of, say, upstream firms for multiple downstream firm partners satisfies a substitutes condition, a sufficient condition for the rank order property for a large number of matching markets it that there be errors facing a social planner in determining the equilibrium in each market. The rank order property allows multiple equilibria to an extent I will discuss. Multiple equilibria is a problem assumed away or even ignored in all previous empirical papers on matching.

After earlier versions of this paper were circulated, Fox and Bajari (2010), Ahlin (2009), Akkus and Hortacsu (2007), Baccara, Imrohoroglu, Wilson and Yariv (2009), Levine (2009), Mindruta (2009), and Yang, Shi and Goldfarb (2009) have conducted empirical work using the matching maximum score estimator I develop here. Their applications are, respectively, matching between bidders and items for sale in a spectrum auction, matching between villagers into risk management groups, mergers between banks after deregulation in the United States, matching between offices and employees with attention paid to several dimensions of social networks, matching between pharmaceutical developers and distributors, matching between individual research team members in the patent development process, and matching between professional athletes and teams with a focus on marketing alliances between players and teams. In addition to my empirical work on automotive suppliers, these disparate applications show the relevance of matching estimation in empirical work in economics, including industrial organization and allied fields such as corporate finance, marketing and strategy.

The paper is organized as follows. Section 2 introduces the deterministic model and Section 3 introduces two rank order properties to make the model stochastic. Section 4 introduces the maximum score estimator and Section 5 provides Monte Carlo evidence. Sections 6–8 comprise the empirical application to automotive suppliers and assemblers. Section 9 concludes.

[1]Dagsvik (2000) provides logit-based methods for studying matching games where other aspects of a relationship than money are also part of the equilibrium matching. Although he does not emphasize it, one-to-one matching games with transferable utility are a special case of his analysis. Matching games with transfers are also related to models of hedonic equilibria, where typically features of the match in addition to price are endogenously determined (Rosen, 1974; Ekeland, Heckman and Nesheim, 2004).

2 Many-to-Many Matching and Pairwise Stability

2.1 Firm Characteristics and Matching Outcomes

This paper studies two-sided, many-to-many matching.[2] The two sides will be **upstream firms** and **downstream firms**. Car parts suppliers are upstream firms and assemblers of cars are downstream firms. Downstream firms match to upstream firms. In the automotive supplier empirical work, a match will actually be a car part, as multiple car parts can be sold from one supplier to one assembler. To outline the model, ignore the complexity of car parts and focus on a match being between a downstream and an upstream firm.

An upstream firm is captured by a vector of characteristics $\tilde{u} \in \tilde{\mathcal{U}}$, where $\tilde{\mathcal{U}} = \mathcal{U} \times (\mathbb{Z}^+ \cup \{\infty\})$ and $\mathcal{U} \subseteq \mathbb{R}^{K_{\text{up}}}$. The first K_{up} elements of $\tilde{u}$ represent **characteristics** that may enter the coming structural revenue functions and the last characteristic is the **quota**, or the number of maximum matches (a positive number or infinity) that the upstream firm can make. For example, $\tilde{u}$ could be $\tilde{u} = \left(u^1, u^2, 3\right)$, where $K_{\text{up}} = 2$, u^1 is a measure of the quality of the products of the firm $\tilde{u}$, u^2 is the firm's past experience and 3 is the maximum number of matches $\tilde{u}$ can make. I also use the notation $u \in \mathcal{U}$ to refer to the characteristics of firm $\tilde{u}$ other than its quota. Likewise, a downstream firm has characteristics $\tilde{d} \in \tilde{\mathcal{D}}$, where $\tilde{\mathcal{D}} = \mathcal{D} \times (\mathbb{Z}^+ \cup \{\infty\})$ and $d \in \mathcal{D}$ for $\mathcal{D} \subseteq \mathbb{R}^{K_{\text{down}}}$. Let the maximum quota of an upstream firm be Q; this can be infinite.

The notation allows for finite numbers of upstream and downstream firms or a continuum (uncountable infinity) of agents. The continuum of agents is important for the asymptotic argument for one large matching market. For the case of a finite number of upstream and downstream firms, we add arbitrary indexes to the definitions of $\tilde{u}$ and $\tilde{d}$ to notationally distinguish two firms with identical characteristics and quotas. For a continuum of agents, notational complexity will require an additional assumption on downstream firms' payoffs, discussed below.

An **outcome** to a matching game with transferable utility is a measure μ on the space $\tilde{\mathcal{U}} \times \left(\tilde{\mathcal{D}} \times \mathbb{R}\right)^Q$, an element of which is a **full partner list** or tuple $\left\langle \tilde{u}, \left(\tilde{d}_1, t_1\right), \ldots, \left(\tilde{d}_N, t_N\right)\right\rangle$ for $N \leq Q$ (N can be infinite if Q is) of the characteristics of one upstream firm $\tilde{u}$, many downstream firms $\tilde{d}_1, \ldots, \tilde{d}_N$, and one possibly negative monetary **transfer** $t_i \in \mathbb{R}$ from each $\tilde{d}$ to $\tilde{u}$. A **full match** is a tuple $\left\langle \tilde{u}, \tilde{d}, t\right\rangle$, where $\tilde{u}$ is the upstream firm involved in the match, $\tilde{d}$ is the downstream firm in the match, and t is the possibly negative monetary transfer from $\tilde{d}$ to $\tilde{u}$. If there are finite numbers of upstream and downstream firms, the outcome measure μ will imply a set $\left\{\left\langle \tilde{u}_1, \tilde{d}_1, t_1\right\rangle, \ldots, \left\langle \tilde{u}_N, \tilde{d}_N, t_N\right\rangle\right\}$ of a finite number N of matches that took place. Note that subscripts as in u_1 refer to firm u_1 and superscripts as in u^1 refer to the first characteristic of firm u. Upstream firm $\tilde{u}$ might have no partners at all in μ, in which case we write that μ gives positive support to the match $\langle \tilde{u}, 0, 0\rangle$. Likewise, the notation $\left\langle 0, \tilde{d}, 0\right\rangle$ refers to an unmatched downstream

[2] Some theoretical results on one-to-one, two-sided matching with transferable utility have been generalized by Kelso and Crawford (1982) for one-to-many matching, Leonard (1983) and Demange, Gale and Sotomayor (1986) for multiple-unit auctions, as well as Sotomayor (1992), Camiña (2006) and Jaume, Massó and Neme (2009) for many-to-many matching with additive separability in payoffs across multiple matches. These models are applications of general equilibrium theory to games with typically finite numbers of agents. The estimator in this paper can be extended to the cases studied by Kovalenkov and Wooders (2003) for one-sided matching, Ostrovsky (2008) for supply chain, multi-sided matching, and Garicano and Rossi-Hansberg (2006) for the one-sided matching of workers into coalitions known as firms with hierarchical production. This paper uses the term "matching game" to encompass a broad class of transferable utility models, including some games where the original theoretical analyses used different names.

firm. For an outcome μ to be feasible in many-to-many matching, the number of downstream firms matched to each $\tilde{u}$ must be less than $\tilde{u}$'s quota and the number of upstream firm matches must be less than each downstream firm's quota.

This paper will work with the case where u and d, but not the quotas and the transfers, are in the data. The next section will discuss econometric unobservables. The notation $\langle u, d, t\rangle$ is a match suppressing the quota, and the notation $\langle u, d\rangle$ is a **physical match**, suppressing quotas and transfers and leaving only observable characteristics. Let M be a set of N physical matches, i.e. $M = \{\langle u_1, d_1\rangle, \ldots, \langle u_N, d_N\rangle\}$, where N can be infinite. Let $\langle u, d_1, \ldots, d_N\rangle$ be a **physical partner list**. Let μ^{A} be the **assignment**, the measure of physical partner lists implied by the measure μ. The assignment will be a superset of the observed data in each market; completely unmatched firms (a potential entrant to making car parts, say) will not be observed in the data. With a finite number of firms in a matching market, another notation for an assignment will be A, the set of observed matches implied by μ^{A}, where again arbitrary firm indices are implicitly used to distinguish two firms with the same characteristics.

Quotas will not enter the payoffs of firms other than as a constraint on the number of matches that they may make. Say u matches with a set D of downstream firms as part of a matching market outcome μ and let $M = \bigcup_{d\in D} \{\langle u, d\rangle\}$. Then, at μ, u gets **profit** $r^{\mathrm{up}}(M) + \sum_{d\in D} t_{\langle u,d\rangle}$, where $r^{\mathrm{up}}(M)$ is the **structural revenue function** of upstream firms as a function of their characteristics and the characteristics of their partners, and $t_{\langle u,d\rangle}$ is the monetary transfer component of the match $\langle u, d, t\rangle$. It is essential that the model allow $r^{\mathrm{up}}(M) \neq \sum_{d\in D} r^{\mathrm{up}}(\{\langle u, d\rangle\})$, or that the structural revenue from multiple matches is not additively separable across downstream firms. Otherwise, the policy question of the gains to a supplier from supplying all of General Motors versus the same set of parts to both GM and a divested former subsidiary Opel could not be answered; the two portfolios of car parts would give the same output. Likewise, let the profit of d for the matches with U, $M = \bigcup_{u\in U} \{\langle u, d\rangle\}$, be $r^{\mathrm{down}}(M) - \sum_{u\in U} t_{\langle u,d\rangle}$. The extra structural revenue from matches of being single or unused quota slots is always 0: $r^{\mathrm{up}}(M \cup \{\langle u, 0\rangle\}) = r^{\mathrm{up}}(M)$ and $r^{\mathrm{down}}(M \cup \{\langle 0, d\rangle\}) = r^{\mathrm{down}}(M)$ for all M.

The case of a continuum of firms is important for the asymptotic argument for one, large matching market. With a continuum, the full partner list notation in the above definition of an outcome μ is not sufficient to describe an unrestricted many-to-many matching situation. In this case, the full partner list notation is sufficient under the additional assumption that downstream firms' payoffs only are additively separable across upstream firms, or $r^{\mathrm{down}}(M) = \sum_{\langle u,d\rangle\in M} r^{\mathrm{down}}(\{\langle u, d\rangle\})$ for M comprised only of matches involving firm d. No such additive separability is imposed for upstream firms.

2.2 Pairwise Stability

The equilibrium concept for both a continuum and a finite number of firms is pairwise stability. The notation $\langle u, d, t\rangle \in \mu$ is a shortened version of writing that there exists a full partner list p in the support of the outcome μ where the full match $\langle u, d, t\rangle$ corresponds to an element of that p. Likewise, $\langle u, d\rangle \in \mu^{\mathrm{A}}$ has a similar meaning for physical matches and assignments.

Definition. An outcome μ will satisfy the equilibrium concept of **pairwise stability** whenever

1. Let $p_1 = \left\langle \tilde{u}_1, \left(\tilde{d}_{1,1}, t_{1,1}\right), \ldots, \left(\tilde{d}_{1,N_1}, t_{1,N_1}\right)\right\rangle$, $p_2 = \left\langle \tilde{u}_2, \left(\tilde{d}_{2,1}, t_{2,1}\right), \ldots, \left(\tilde{d}_{2,N_2}, t_{2,N_2}\right)\right\rangle$, $d_1 \in \{d_{1,1}, \ldots, d_{1,N_1}\}$, $d_2 \in \{d_{2,1}, \ldots, d_{2,N_2}\}$, $M_{u_1} = \{\langle u_1, d_{1,1}\rangle, \ldots, \langle u_1, d_{1,N}\rangle\}$ and $M_{d_2} = \left\{\langle u, d_2\rangle \in \mu^A\right\}$. The following inequality holds for all full partner lists $p_1 \in \mu$ and $p_2 \in \mu$:

$$r^{\text{up}}\left(M_{u_1}\right) + t_{\langle u_1, d_1\rangle} \geq r^{\text{up}}\left(\left(M_{u_1} \setminus \{\langle u_1, d_1\rangle\}\right) \cup \{\langle u_1, d_2\rangle\}\right) + \tilde{t}_{\langle u_1, d_2\rangle}, \tag{1}$$

where $\tilde{t}_{\langle u_1, d_2\rangle} \equiv r^{\text{down}}\left(\left(M_{d_2} \setminus \{\langle u_2,, d_2\rangle\}\right) \cup \{\langle u_1, d_2\rangle\}\right) - \left(r^{\text{down}}\left(M_{d_2}\right) - t_{\langle u_2, d_2\rangle}\right)$.

2. The inequality (1) holds if either or both of the existing matches represent a free quota slot, namely $\langle u_1, d_1\rangle = \langle u_1, 0\rangle$ or $\langle u_2, d_2\rangle = \langle 0, d_2\rangle$. In this case, in (1) set the transfers corresponding to the free quota slots, $t_{\langle u_1, d_1\rangle}$ or $t_{\langle u_2, d_2\rangle}$, equal to 0.

3. For all $\langle u, d, t\rangle \in \mu$ for any p, where $M_u = \{\langle u, d_1\rangle, \ldots, \langle u, d_N\rangle\}$ and $d \in \{d_1, \ldots, d_N\}$,

$$r^{\text{up}}\left(M_u\right) + t_{\langle u, d\rangle} \geq r^{\text{up}}\left(M_u \setminus \{\langle u, d\rangle\}\right).$$

4. For all $\langle u, d, t\rangle \in \mu$ for any p, where $M_d = \{\langle u_1, d\rangle, \ldots, \langle u_N, d\rangle\}$ and $u \in \{u_1, \ldots, u_N\}$,

$$r^{\text{down}}\left(M_d\right) - t_{\langle u, d\rangle} \geq r^{\text{down}}\left(M_d \setminus \{\langle u, d\rangle\}\right).$$

Part 1 of the definition of pairwise stability says that u_1 prefers its matched downstream firm d_1 instead the alternative d_2 at the transfer $\tilde{t}_{\langle u_1, d_2\rangle}$ that makes d_2 switch to sourcing its supplies from u_1 instead of its equilibrium partner u_2. Because of transferable utility, u_1 can always cut its price and attract d_2's business; at a pairwise stable equilibrium, u_1 would lower its profit from doing so if the new business supplanted the match with d_1. Part 1 is the component of the definition of pairwise stability that estimation is indirectly based on.

Part 2 deals with firms with free quota slots, including completely unmatched firms, not adding new matches or exchanging old matches for new matches. Parts 3 and 4 deal with matched firms not profiting by unilaterally dropping a relationship and becoming unmatched. These are individual rationality conditions: all matches must give an incremental positive surplus. Parts 2–4 compare being matched to unmatched, and so implementing the restrictions from Parts 2–4 requires data on unmatched firms. A person being single or unmarried is often found in marriage data. The notion that a car parts supplier in an upstream–downstream market would have a free quota slot or be a potential entrant is a modeling abstraction. It is often hard to find data on quotas and potential entrants.

2.3 Sum of Revenues Inequalities

I wish to work with an implication of pairwise stability that does not involve data on transfers. Substituting the expression for $\tilde{t}_{\langle u_1, d_2\rangle}$ into (1) gives

$$\begin{aligned} &r^{\text{up}}\left(M_{u_1}\right) + t_{\langle u_1, d_1\rangle} + r^{\text{down}}\left(M_{d_2}\right) \geq \\ &\quad r^{\text{up}}\left(\left(M_{u_1} \setminus \{\langle u_1, d_1\rangle\}\right) \cup \{\langle u_1, d_2\rangle\}\right) + r^{\text{down}}\left(\left(M_{d_2} \setminus \{\langle u_2,, d_2\rangle\}\right) \cup \{\langle u_1, d_2\rangle\}\right) + t_{\langle u_2, d_2\rangle}. \end{aligned} \tag{2}$$

A symmetric inequality holds for u_2 not wanting to replace d_2 with d_1,

$$\begin{aligned}
&r^{\mathrm{up}}\left(M_{u_2}\right)+t_{\langle u_2,d_2\rangle}+r^{\mathrm{down}}\left(M_{d_1}\right)\geq \\
&\qquad r^{\mathrm{up}}\left(\left(M_{u_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)+r^{\mathrm{down}}\left(\left(M_{d_1}\setminus\{\langle u_1,,d_1\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)+t_{\langle u_1,d_1\rangle},
\end{aligned}\tag{3}$$

where the notation is analogous to that in part 1 of the definition of pairwise stability. Adding (2) and (3) cancels the transfers and gives

$$\begin{aligned}
&r^{\mathrm{up}}\left(M_{u_1}\right)+r^{\mathrm{down}}\left(M_{d_1}\right)+r^{\mathrm{up}}\left(M_{u_2}\right)+r^{\mathrm{down}}\left(M_{d_2}\right)\geq \\
&\qquad r^{\mathrm{up}}\left(\left(M_{u_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right)+r^{\mathrm{down}}\left(\left(M_{d_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)+ \\
&\qquad\qquad r^{\mathrm{up}}\left(\left(M_{u_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)+r^{\mathrm{down}}\left(\left(M_{d_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right).
\end{aligned}\tag{4}$$

The inequality (4) is called a **sum of revenues inequality** because it compares the sum of structural revenues of two upstream firms and two downstream firms, before and after an exchange of one downstream firm each between two upstream firms. Sum of revenues inequalities will form the basis for the maximum score estimation approach.

2.4 Equilibrium Existence and Uniqueness

A pairwise stable equilibrium is not guaranteed to exist in many-to-many matching games. Nor is a pairwise stable equilibrium guaranteed to be unique. In my computational experience with simple many-to-many matching games, multiplicity is a more common occurrence than non-uniqueness. For the parallel case of many-to-one, non-transferable utility matching games, Kojima, Pathak and Roth (2010) find empirically and theoretically that the lack of a pairwise stable outcome is often not a major concern.

If the revenue functions of upstream and downstream firms satisfy a condition known as **substitutes**, then a pairwise stable outcome will be guaranteed to exist and will be equivalent to fully stable outcome where any coalition of firms can consider deviating at once (Milgrom, 2000; Hatfield and Milgrom, 2005; Hatfield and Kominers, 2010). As the entire coalition of firms can deviate, in transferable utility games a fully stable outcome will maximize the sum of revenues across entire physical assignments A or μ^{A},

$$\sum_u r^{\mathrm{up}}\left(M_u^A\right)+\sum_d r^{\mathrm{down}}\left(M_d^A\right),$$

where M_d^A is the set of upstream firms matched to downstream firm d at A and the sums imply a finite number of total firms, for simplicity. Then under substitutable preferences, a pairwise stable assignment can be computed by a linear programming problem. Further, if the characteristics u and d have continuous supports with no atoms, the probability that any two assignments both maximize the sum of revenues will be 0. So substitutes is a useful condition: it ensures existence, it gives uniqueness with probability 1, and it provides a computationally simple algorithm to compute a pairwise stable outcome. Unfortunately, the substitutes condition will not apply to automotive suppliers, as selling two parts to General Motors may give more structural revenue than selling one car part to General

Motors and another to a divested Opel. How existence and multiplicity affect estimation will be discussed more in the next section.

3 The Rank Order Properties

This section allows each physical partner list or each overall assignment to have a positive probability, which makes the previously deterministic matching game stochastic. This section introduces two so-called rank order properties, corresponding to different asymptotic arguments for the consistency of the eventual estimator. The first argument involves one large matching market and the second argument involves many independent matching markets.

Notationally, a stochastic structure $S \in \mathcal{S}$ will index distributions of unobservables. Each rank order property is imposed as a primitive, but sufficient conditions on classes $\mathcal{S}$ of stochastic structures will be given as assumptions on distributions of heterogeneity that imply each rank order property, for special cases. This approach to motivating the consistency of the estimator will be helpful because of the computational simplicity of the estimator, which I will discuss below.

3.1 Rank Order Property for One Large Matching Market

A researcher may have data on one large matching market. For example, Choo and Siow (2006) study the US marriage market and Fox and Bajari (2010) study a large FCC spectrum auction. In these papers, the asymptotic fiction is that the observed matches in the data correspond to a finite set of observations from a true matching game with a continuum of agents and matches. Keep in mind that any asymptotic argument is designed to mimic the finite sample properties of an estimator rather than to describe how additional entry would affect an upstream downstream market.

As discussed previously, when the true matching market is a continuum, it is notationally necessary to impose additive separability in downstream firms' structural revenue functions, $r^{\text{down}}(M) = \sum_{\langle u,d\rangle \in M} r^{\text{down}}(\{\langle u,d\rangle\})$. Under this restriction, cancelling terms that are the same on both sides of the sum of revenues inequality (4) gives

$$\begin{aligned} r^{\text{up}}(M_{u_1}) + r^{\text{down}}(\{\langle u_1,d_1\rangle\}) + r^{\text{up}}(M_{u_2}) + r^{\text{down}}(\{\langle u_2,d_2\rangle\}) \geq \\ r^{\text{up}}((M_{u_1} \setminus \{\langle u_1,d_1\rangle\}) \cup \{\langle u_1,d_2\rangle\}) + r^{\text{down}}(\{\langle u_2,d_1\rangle\}) \\ + r^{\text{up}}((M_{u_2} \setminus \{\langle u_2,d_2\rangle\}) \cup \{\langle u_2,d_1\rangle\}) + r^{\text{down}}(\langle u_1,d_2\rangle), \end{aligned} \quad (5)$$

which does not require knowledge of the other matches of downstream firms d_1 and d_2.

Further assume that the assignment measure μ^{A}, from the overall outcome measure μ, admits a density function g over physical partner lists $\langle u, d_1, \ldots, d_N\rangle$. The density function can be with respect to the counting measure for characteristics in u or d that are discrete. To emphasize that the assignment density g is an equilibrium (although aggregately deterministic) outcome to a matching game with a continuum of agents, I write $g^{r^{\text{up}},r^{\text{down}},S}$, where the superscripts refer to three unknown functions: the two structural revenue functions and the distribution of unobservables. The density $g^{r^{\text{up}},r^{\text{down}},S}$ itself is not stochastic, but each upstream firm u's list of partners $(d_1, \ldots, d_N)$ is a random draw from the conditional density of $(d_1, \ldots, d_N)$ given u.

Property 3.1. *Let r^{up}, r^{down} and S be given. Let $p_1 = \langle u_1, d_{1,1}, \ldots, d_{1,N_1} \rangle$, $D_1 = \{d_{1,1}, \ldots, d_{1,N_1}\}$, $M_{u_1} = \{\langle u_1, d_{1,1} \rangle, \ldots, \langle u_1, d_{1,N_1} \rangle\}$, $d_1 \in D_1$, $p_2 = \langle u_2, d_{2,1}, \ldots, d_{2,N_2} \rangle$, $D_2 = \{d_{2,1}, \ldots, d_{2,N_2}\}$, $M_{u_2} = \{\langle u_2, d_{2,1} \rangle, \ldots, \langle u_2, d_{2,N_2} \rangle\}$, and $d_2 \in D_2$. Also let p_3 be the physical partner list formed from $(M_{u_1} \setminus \{\langle u_1, d_1 \rangle\}) \cup \{\langle u_1, d_2 \rangle\}$ and p_4 be the physical partner list formed from $(M_{u_2} \setminus \{\langle u_2, d_2 \rangle\}) \cup \{\langle u_2, d_1 \rangle\}$.*

The ***rank order property for one large market*** *states that the sum of revenues inequality* (5) *holds if and only if*

$$g^{r^{\text{up}}, r^{\text{down}}, S}(p_1) \cdot g^{r^{\text{up}}, r^{\text{down}}, S}(p_2) \geq g^{r^{\text{up}}, r^{\text{down}}, S}(p_3) \cdot g^{r^{\text{up}}, r^{\text{down}}, S}(p_4). \tag{6}$$

The rank order property for one large market allows the sum of revenues inequality (5) to hold for some sets of four physical partner lists, two on the left and two on the right, and to be violated for other partner lists. An inequality might be violated because of unobservables to the econometrician. However, pairs of two physical partner lists such that the sum of deterministic revenues on the left side of (5) exceed those on the right side are more likely to be jointly observed that the pairs of two physical partner lists on the right side. This rank order property is a natural extension of the deterministic implications of pairwise stability, the sum of revenues inequality (5), to the case of an econometric model where all physical partner lists may be in the support of the data generating process.

The rank order property for one large market assumes that an equilibrium exists. Multiple equilibria do not pose a problem because the equilibrium in the data is conditioned on.

3.1.1 A Sufficient Condition for One-to-One Matching

A sufficient condition for the rank order property for one large market in the case of many-to-many matching is not known, in part because there is no existing theoretical or empirical literature on many-to-many matching games with a continuum of agents and econometric errors. The previous paper on estimating one-to-one matching games of transferable utility (marriage) is the logit based model of Choo and Siow (2006). Choo and Siow use a model where each u and d is a set of characteristics with finite support, there is an infinite number of firms, and firms have heterogeneous preferences over the types (the values of u and d) of potential partners. Assume that each upstream firm and each downstream firm can make at most one match (a quota of 1). At the outcome μ, the profit of upstream firm i with characteristics u for downstream firm d is $r^{\text{up}}(\{\langle u, d \rangle\}) + \psi_{i,d} + t_{\langle u,d \rangle}$, where $\psi_{i,d}$ has the type I extreme value distribution familiar from the literature on the multinomial logit (McFadden, 1973). Likewise, the utility of downstream firm j with characteristics d for upstream firm u is $r^{\text{down}}(\{\langle u, d \rangle\}) + \psi_{j,u} - t_{\langle u,d \rangle}$. The implied logit choice probabilities give a set of demand equations for upstream firms and for downstream firms for matches of each type u and d, and the equilibrium transfers $t_{\langle u,d \rangle}$ equate the demand for each match type from both sides of the market.

Proposition. *The Choo and Siow (2006) matching model satisfies the rank order property for one large market.*

Thus, the rank order property for one large market is strictly more general than the only previous paper on estimating transferable utility matching games. The rank order property does not impose the parametric type I extreme value errors, but is consistent in their presence.

Proof. In one-to-one matching, the physical partner lists are $p_1 = \langle u_1, d_1 \rangle$, $p_2 = \langle u_2, d_2 \rangle$, $p_3 = \langle u_1, d_2 \rangle$ and $p_4 = \langle u_2, d_1 \rangle$. Also, Choo and Siow allow only discrete characteristics, so the density $g^{r^{\text{up}}, r^{\text{down}}, S}$ is also a mass function. Rearranging equation (10) in Choo and Siow gives, in my notation,

$$g^{r^{\text{up}}, r^{\text{down}}, S}(p_1) =$$
$$\exp\left(\frac{1}{2}\left(r^{\text{up}}(\{\langle u_1, d_1 \rangle\}) + r^{\text{down}}(\{\langle u_1, d_1 \rangle\}) + \log g^{r^{\text{up}}, r^{\text{down}}, S}(\langle u_1, 0 \rangle) + \log g^{r^{\text{up}}, r^{\text{down}}, S}(\langle 0, d_1 \rangle)\right)\right),$$

where the last two terms refer to the frequencies of unmatched firms of types u_1 and d_1. Substituting the Choo and Siow equilibrium match or partner list probabilities for p_1–p_4 into inequality (6), simplifying, taking logarithms of both sides and cancelling the fractions of each type that are single, which are the same on both sides of (6), gives

$$\begin{aligned} &r^{\text{up}}(\{\langle u_1, d_1 \rangle\}) + r^{\text{down}}(\{\langle u_1, d_1 \rangle\}) + r^{\text{up}}(\{\langle u_2, d_2 \rangle\}) + r^{\text{down}}(\{\langle u_2, d_2 \rangle\}) \geq \\ &\qquad r^{\text{up}}(\{\langle u_1, d_2 \rangle\}) + r^{\text{down}}(\{\langle u_1, d_2 \rangle\}) + r^{\text{up}}(\{\langle u_2, d_1 \rangle\}) + r^{\text{down}}(\{\langle u_2, d_1 \rangle\}). \end{aligned}$$

By inspection, this is the appropriate simplification of the sum of revenues inequality (5) for one-to-one matching. □

3.2 The Rank Order Property for Many Independent Matching Markets

Another data generating process is to observe data from many independent matching markets with the same structural revenue functions. By independent matching markets, I mean that upstream firms in one market cannot match with downstream firms in another market. Because of the history of the automotive supplier industry, where particular firms often have manufactured the same types of car parts since the early twentieth century, I will model each car component category as a separate matching market.

Each matching market will have a finite set of both upstream and downstream firms, although the number of firms can differ across markets. Within each market, we will observe the set of matches or A, the assignment. Only the portion of the assignment pertaining to firms in realized matches may be observed, as I do not have data on potential entrants who in equilibrium supply no car parts. I will not introduce new notation to reflect this missing data; A will represent matches with potential entrants discarded. There is no need to impose additive separability in the structural revenue function of downstream firms.

Let $\mu^{\tilde{d}}$ and $\mu^{\tilde{u}}$ be the measures of the characteristics of upstream and downstream firms, including quotas, implied by the measure μ. These are exogenous characteristics in a matching game. Let the measure $\nu\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$ describe how the exogenous characteristics of firms vary across matching markets. Let ψ describe a vector of econometric unobservables and let $S\left(\psi \mid \mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$ be the distribution of ψ conditional on the measures of the characteristics of upstream and downstream firms. Details on ψ

will be given below. The data generating process will imply a frequency, a density ρ over both discrete and continuous characteristics, of observing each assignment A, where the assignment contains only non-single matches and quotas are not observed,

$$\rho_{r^{\text{up}},r^{\text{down}},S,\nu}(A) \propto \int 1\left[A\,\text{pairwise stable} \mid \mu^{\tilde{d}},\mu^{\tilde{u}},\psi;\ r^{\text{up}},r^{\text{down}}\right]\cdot$$
$$1\left[A\,\text{selected} \mid A\,\text{pairwise stable},\mu^{\tilde{d}},\mu^{\tilde{u}},\psi;\ r^{\text{up}},r^{\text{down}},S\right] dS\left(\psi \mid \mu^{\tilde{d}},\mu^{\tilde{u}}\right) d\nu\left(\mu^{\tilde{d}},\mu^{\tilde{u}}\right),$$

where the symbol $\propto$ refers to proportional to, to emphasize that the portion of the density that is written may not integrate to 1. There are four terms in the integrand: an indicator for whether A is the assignment portion of a pairwise stable outcome given the observed firm characteristics and unobservables, an indicator for whether A is the selected assignment in the case where multiple assignments may be parts of pairwise stable outcomes, the distribution of the unobservables, and the distribution of the (mostly) observable firm characteristics.[3] The portion of the density that is written may not integrate to 1 also in the case where a pairwise stable matching does not exist for some $\left(\mu^d,\mu^s,\psi\right)$, although I argued above non-existence happens infrequently in simulations.[4]

The theory of matching games is more informative about pairwise stability than equilibrium assignment selection rules. Therefore let the density

$$\Upsilon_{r^{\text{up}},r^{\text{down}},S,v}(A) \propto \int 1\left[A\,\text{pairwise stable} \mid \mu^{\tilde{d}},\mu^{\tilde{u}},\psi;\ r^{\text{up}},r^{\text{down}}\right] dS\left(\psi \mid \mu^{\tilde{d}},\mu^{\tilde{u}}\right) d\nu\left(\mu^{\tilde{d}},\mu^{\tilde{u}}\right).$$

Property 3.2. *Let r^{up}, r^{down}, S and ν be given. Let A_1 be an assignment and let*

$$A_2 = \left(A_1 \setminus \{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\}\right) \cup \{\langle u_1,d_2\rangle,\langle u_2,d_1\rangle\}$$

be the assignment formed by removing the matches $\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\} \subseteq A_1$ and replacing them with the exchange of partners $\{\langle u_1,d_2\rangle,\langle u_2,d_1\rangle\}$. Also, let $M_{u_1} \subseteq A_1$, $M_{u_2} \subseteq A_1$, $M_{d_1} \subseteq A_1$ and $M_{d_2} \subseteq A_1$ be the matches for the respective firms under assignment A_1.

*The **rank order property for many markets** states that the sum of revenues inequality* (4) *holds if and only if the following two conditions jointly hold:*

1. $\Upsilon_{r^{\text{up}},r^{\text{down}},S,\nu}(A_1) \geq \Upsilon_{r^{\text{up}},r^{\text{down}},S,\nu}(A_2)$ *and*
2. $\rho_{r^{\text{up}},r^{\text{down}},S,\nu}(A_1) \geq \rho_{r^{\text{up}},r^{\text{down}},S,\nu}(A_2)$ *if and only if* $\Upsilon_{r^{\text{up}},r^{\text{down}},S,\nu}(A_1) \geq \Upsilon_{r^{\text{up}},r^{\text{down}},S,\nu}(A_2)$.

More succinctly, the rank order property for many markets implies that the sum of revenues inequality (4) holds if and only if

$$\rho_{r^{\text{up}},r^{\text{down}},S,\nu}(A_1) \geq \rho_{r^{\text{up}},r^{\text{down}},S,\nu}(A_2). \tag{7}$$

Keep in mind that r^{up}, r^{down}, S and ν are fixed; the rank order property is a property of the stochastic structure of the model and the equilibrium assignment selection rule. The rank order

[3] The phrase "A pairwise stable" is shorthand for A being the assignment portion, without potential entrants, of a pairwise stable outcome.

[4] Non-existence occurs also in Nash games when attention is restricted to pure strategies.

property compares two nearly identical assignments that differ only because upstream firms u_1 and u_2 exchange one downstream firm partner each. Neither A_1 or A_2 may be the assignment portion of a pairwise stable outcome to the matching model without error terms. But A_1 might dominate A_2 in the deterministic model in that at least two firms in A_2 (either u_1 and d_2 or u_2 and d_1) would prefer to match with each other instead of their assigned partners, leading to A_1. The rank order property states that A_1 is more likely to be observed than A_2.

Unmatched firms are not necessarily recorded in the data and neither are quotas of firms. The rank order property for many markets does not require data on either; the same set of firms can be unmatched when the set of realized matches are either A_1 or A_2. Likewise, the number of matches that each firm has is the same in A_1 and A_2. If A_1 does not violate quotas for some $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$, A_2 will not violate quotas either for that $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$. Therefore, quotas will not affect the rank ordering of A_1 and A_2.

The equilibrium assignment selection rule component of the rank order property for many markets preserves the rank ordering of pairwise stability: assignments that are more likely to be pairwise stable are more likely to occur. The rank order property will give a simple maximum score estimator, regardless of the number of pairwise stable assignments for each realization of $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}, \psi\right)$.[5]

3.2.1 A Sufficient Condition for Many-to-Many Matching Under Substitutes

There is a unique equilibrium assignment with probability 1 if one is willing to assume that the structural revenue functions of upstream firms for multiple downstream firms and of downstream firms for multiple upstream firms both exhibit the substitutes condition. Under substitutes, the equilibrium assignment rule does not enter the data generating process and $\Upsilon_{r^{\text{up}}, r^{\text{down}}, S, \nu}(A) = \rho_{r^{\text{up}}, r^{\text{down}}, S, \nu}(A)$. Further, the equilibrium assignment maximizes the sum of structural revenues in the economy.

A sufficient condition for the rank order property with many markets and many-to-many matching under substitutes follows. Let the data generating process be that the social planner picks the assignment A to maximize $\sum_u r^{\text{up}}\left(M_u^A\right) + \sum_d r^{\text{up}}\left(M_d^A\right) + \psi_A$, where $M_u^A \subseteq A$ is the set of matches involving upstream firm u in the assignment A and where ψ_A is an error term for assignment A that enters the social planner's payoff for assignment A. Let $\psi = (\psi_A)$ be the vector of assignment level errors for all feasible assignments, given a realization of $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$.

Proposition 3.1. *Let the payoff to assignment A to a social planner be $\sum_u r^{\text{up}}\left(M_u^A\right) + \sum_d r^{\text{up}}\left(M_d^A\right) + \psi_A$ and let the distribution $S\left(\psi \mid \mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$ be such that ψ is an exchangeable random vector for each realization of $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$. Then the rank order property with many matching markets is satisfied.*

This lemma was proved in Goeree, Holt and Palfrey (2005) and is a generalization of a result in Manski (1975).[6] The proposition casts the choice of assignment A as a single agent discrete choice problem. Assignments with higher deterministic payoffs $\sum_u r^{\text{up}}\left(M_u^A\right) + \sum_d r^{\text{up}}\left(M_d^A\right)$ will occur more

[5] The literature on estimating parametric Nash games, a non-nested class with matching games, presents strategies with perhaps fewer assumptions but higher computational demands in estimation for dealing with multiple equilibria. See Bajari, Hong and Ryan (2010) and Ciliberto and Tamer (2009).

[6] An exchangeable random vector $(y_1, \ldots, y_n)$ has the same distribution as $(\pi y_1, \ldots, \pi y_n)$ for any permutation π. The proof in Goeree, Holt and Palfrey conditions on $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$. As the property holds for each $\left(\mu^{\tilde{d}}, \mu^{\tilde{u}}\right)$, it holds for the unconditional probabilities $\rho_{r^{\text{up}}, r^{\text{down}}, S, \nu}(A)$.

often if ψ is exchangeable. One could then view exchangeability of econometric unobservables as a structural assumption on the equilibrium-assignment selection process. Adding errors to a deterministic model is similar to the quantile-response-equilibrium method of perturbing behavior (Goeree et al.). The social planning problem is a structural assumption that does exactly generalize the intuition from the informal empirical literature following the work of Becker (1973) on marriage that assignments that give higher output from observable characteristics are more likely to occur.

The sufficient conditions for both rank order properties do not allow for the firm- but not match-specific unobservables empirically found to be important in Ackerberg and Botticini (2002). I am investigating firm-specific unobservables in other work; their presence will not lead to a computationally simple, maximum score estimator.

4 The Maximum Score Estimator

I now discuss how maximum score can form the basis for a practical estimator. The maximum score estimator avoids a computational curse of dimensionality by not performing integrals or nested computations of equilibrium assignments. Further, all inequalities do not need to be included with probability 1 to maintain the consistency of the estimator. Maximum score estimation was introduced by Manski (1975, 1985) for the single-agent model.

Whatever the asymptotic argument may be, in a finite sample the dataset records a finite number of matches in the assignment set A_h for markets $h = 1, \ldots, H$. It may be that $H = 1$ but there is a lot of information in a single market, or it may be that $H > 1$. I assume that the A_h are i.i.d. across markets when $H > 1$.

The estimator is semiparametric in that S will not be specified up to a finite vector of parameters. Indeed, following Manski (1975) and later work on maximum score estimation of the single agent choice model, S will not be estimated. The structural revenue functions $r^{\text{up}}_{\beta^{\text{up}}}$ and $r^{\text{down}}_{\beta^{\text{down}}}$ will be specified up to a finite vector of parameters $\beta = \left(\beta^{\text{up}}, \beta^{\text{down}}\right)$. The parameter vector β is the object of estimation.

4.1 Revenue Functions That Are Linear in Parameters

For simplicity, I restrict attention to structural revenue functions that are linear in the estimable parameters. It is not conceptually difficult to weaken the linear in parameters restriction as in Matzkin (1993) for the polychotomous choice model and my own nonparametric identification results for matching games in Fox (2010).

Recall that firms are indexed by their characteristics u or d and that M is a set of matches. For upstream firms, $r^{\text{up}}_{\beta^{\text{up}}}(M) = Z^{\text{up}}(M)' \beta^{\text{up}}$, where $Z^{\text{up}}(M)$ is a vector-valued function of M. Likewise, $r^{\text{down}}_{\beta^{\text{down}}}(M) = Z^{\text{down}}(M)' \beta^{\text{down}}$. In empirical work, the researcher chooses $Z^{\text{up}}(M)$ to capture aspects of the characteristics of the downstream and upstream firms matched in M that will contribute to an upstream firm's revenue. The choice of the regressors in $Z^{\text{up}}(M)$ is guided by the context of the empirical investigation, most importantly the institutional details of the industry under study.

Under this choice of functional forms, the sum of revenues inequality (4) becomes

$$\begin{aligned}
&Z^{\text{up}}\left(M_{u_1}\right)'\beta^{\text{up}}+Z^{\text{down}}\left(M_{d_1}\right)'\beta^{\text{down}}+Z^{\text{up}}\left(M_{u_2}\right)'\beta^{\text{up}}+Z^{\text{down}}\left(M_{d_2}\right)'\beta^{\text{down}}\geq\\
&\quad Z^{\text{up}}\left(\left(M_{u_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right)'\beta^{\text{up}}+Z^{\text{down}}\left(\left(M_{d_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)'\beta^{\text{down}}\\
&\quad+Z^{\text{up}}\left(\left(M_{u_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)'\beta^{\text{up}}+Z^{\text{down}}\left(\left(M_{d_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right)'\beta^{\text{down}}. \quad (8)
\end{aligned}$$

This can be simplified by defining X_{u_1,u_2,d_1,d_2} to be one long vector composed of the elements of the vectors $X^{\text{up}}_{u_1,u_2,d_1,d_2}$ and $X^{\text{down}}_{u_1,u_2,d_1,d_2}$, where

$$X^{\text{up}}_{u_1,u_2,d_1,d_2}=Z^{\text{up}}\left(M_{u_1}\right)+Z^{\text{up}}\left(M_{u_2}\right)-Z^{\text{up}}\left(\left(M_{u_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right)-Z^{\text{up}}\left(\left(M_{u_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)$$

$$\begin{aligned}
X^{\text{down}}_{u_1,u_2,d_1,d_2}&=Z^{\text{down}}\left(M_{d_1}\right)+Z^{\text{down}}\left(M_{d_2}\right)-\\
&Z^{\text{down}}\left(\left(M_{d_1}\setminus\{\langle u_1,d_1\rangle\}\right)\cup\{\langle u_2,d_1\rangle\}\right)-Z^{\text{down}}\left(\left(M_{d_2}\setminus\{\langle u_2,d_2\rangle\}\right)\cup\{\langle u_1,d_2\rangle\}\right).
\end{aligned}$$

With this notation, the inequality (8) simplifies to $X_{u_1,u_2,d_1,d_2}{}'\beta\geq 0$.

There are two special issues to highlight for identification. The first issue is the inability to identify a parameter on a firm characteristic that is not interacted with the characteristics of any other firm. The regressors $Z^{\text{up}}(M)$ and $Z^{\text{down}}(M)$ must only capture interactions of the characteristics between two or more firms. If $M=\{\langle u,d\rangle\}$, the choice of $Z^{\text{up}}(M)=\left(u^1,u^2,d^1,d^2\right)$ for four scalar firm characteristics, two for $u=\left(u^1,u^2\right)$ and two for $d=\left(d^1,d^2\right)$, will not lead to identification of β^{up}. The same firms appear on the left and right sides of the sum of revenues inequality (8) and so additive terms that are not interactions between the characteristics of different firms will be the same on the left as on the right, and will cancel out of the inequality. In notation, $X^{\text{up}}_{u_1,u_2,d_1,d_2}=0$ for all pairs of matches $\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\}$. In a matching game with transferable utility, characteristics of one firm that are not interacted with those of another firm are priced out in the pairwise stable outcome and do not affect the stable assignment, at least among the set of firms that do not have unused quota.

The second special issue for identification involves the ability to separately identify β^{up} and β^{down}. Separate identification of β^{up} and β^{down} requires that the characteristics in one of either $Z^{\text{up}}(M)$ or $Z^{\text{down}}(M)$ involve the interactions of, respectively, two or more downstream firms with one upstream firm or two or more upstream firms with one downstream firm. If the sum of values inequality (8) is indexed by $\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\}$, the downstream firm characteristics d_3 from the match $\langle u_1,d_3\rangle$ or the upstream firm characteristics u_4 from the match $\langle u_4,d_1\rangle$ provide exclusion restrictions that allow us to learn how much of the structural revenue from the characteristics in u_1 and d_1 occurs to u_1 and how much occurs to d_1. By exclusion restriction, I am saying there are matching arrangements where u_1 is matched with d_3 and u_2 is not, so d_3 enters the inequality only through the revenue of u_1. If the interaction between the characteristics u_1, d_1 and d_3 is important, than we attribute the revenue to the upstream firm u_1 and if the interaction between the characteristics u_1, d_1 and u_4 is important, we attribute the revenue to the downstream firm d_1. If one element of both the vectors $Z^{\text{up}}(M)$ and $Z^{\text{down}}(M)$ is a simple interaction between two scalar characteristics, $u^1\cdot d^1$, we identify the sum of the corresponding elements of β^{up} and β^{down}. We cannot learn how much of the revenue accrues to upstream and to downstream firms, as the characteristic $u^1\cdot d^1$ in $Z^{\text{up}}(M)$ is linearly dependent with itself in $Z^{\text{down}}(M)$ in the inequality (8).

A special case of many-to-many matching is one-to-many matching. In that case, there are no exclusion restrictions from the characteristics of additional matches and elements of $Z^{\text{up}}(M)$ and $Z^{\text{down}}(M)$ will, for example, be of the form $u^1 \cdot d^1$. In this case, without imposing that some interactions of characteristics are not valued by either upstream or downstream firms, one identifies the sum of β^{up} and β^{down}. Fox (2010) calls the sum of the revenues of the two sides of the market the production function for a match, and explores its nonparametric identification.[7] The ability, in many-to-many matching, to separately identify the revenue functions of both sides of the market is new to this paper.

4.2 The Matching Maximum Score Estimator

There are a variety of inequalities that could be included for each market. Given A_h for market h, let I_h be the inequalities that the econometrician includes for market h. An inequality in I_h is indexed by the matches $\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\} \subseteq A_h$ on the left side of the sum of revenues inequality (4). Not all inequalities may be included for computational and for data availability reasons. For example, data on unmatched firms may not be available. The maximum score estimator is any parameter vector $\hat{\beta}_H$ that maximizes

$$Q_H(\beta) = \frac{1}{H} \sum_{h \in H} \sum_{\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\} \in I_h} 1\left[X_{u_1,u_2,d_1,d_2}{}'\beta \geq 0\right]. \tag{9}$$

Evaluating $Q_H(\beta)$ is computationally simple: there is no nested equilibrium computation to a matching game, as say Pakes (1986) and Rust (1987) proposed for dynamic programming problems. Another key idea behind the computational simplicity of maximum score estimation is that there are no econometric unobservable terms and hence no integrals in (9). Because of this, not all inequalities will be satisfied, even at the maximizer $\hat{\beta}_H$ and even at the probability limit of the objective function.[8]

Manski and Thompson (1986) and Pinkse (1993) present optimization algorithms for the maximum score objective function where the parameters enter linearly into the utility function. In the empirical work, I numerically maximize the maximum score objective function using the global optimization routine known as differential optimization (Storn and Price, 1997). Visually, the objective function may look rather smooth when viewed from far away, when there is a large number of inequalities. The estimator is point identified when the number of markets grows large; the limiting objective function is smooth. In a finite sample, researchers must take care to run their optimizer many times in order to ensure that they have found the global optimum. Such care should be taken for most optimization problems; this concern is not specific to maximum score.

[7]These semiparametric identification arguments parallel the nonparametric identification arguments in Fox (2010), who argues that, say, the nonparametric analog of identifying the elements of the sum $\beta^{\text{up}} + \beta^{\text{down}}$ corresponding to $u^1 \cdot d^1$ is identifying the cross-partial derivative of the production function with respect to u^1 and d^1. Another result in Fox (2010) is that vertical characteristics and horizontal characteristics can be distinguished in production: the functions $-\left(u^1 - d^1\right)^2$ and $u^1 \cdot d^1$ can be distinguished. For firm-specific characteristics, this result relies on the individual rationality decision to remain unmatched.

[8]This distinguishes maximum score from a moment inequality estimator (Pakes, Porter, Ho and Ishii, 2006).

4.3 Choosing Inequalities

The set of inequalities I_h included in estimation for market h does not need to include all theoretically valid inequalities. If all inequalities were included, the estimator would suffer from a computational curse of dimensionality in the number of firms in a matching market, as the number of valid inequalities grows rapidly with the number of firms in the market. In the car parts empirical work, one automotive component category has 3.1 million possible inequalities. Luckily, inequalities only need to be included with some positive probability for the estimator to be consistent as $H \to \infty$.[9] This means researchers can sample from the set of theoretically valid inequalities. Let $W(A)$ be this set of theoretically valid sum of revenues inequalities of the form $\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\}$ given the assignment A. Let $C(\langle u_1, d_1\rangle, \langle u_2, d_2\rangle)$ be the probability that a researcher includes an inequality when $\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\} \in W(A)$. Hence, $C(\langle u_1, d_2\rangle, \langle u_2, d_1\rangle)$ is the probability of sampling $\{\langle u_1, d_2\rangle, \langle u_2, d_1\rangle\}$ when $\{\langle u_1, d_2\rangle, \langle u_2, d_1\rangle\} \in W(A_2)$ for some other assignment A_2.

Assumption. *For all* $\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\} \in W(A)$,

1. $C(\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\}) > 0$.

2. $C(\langle u_1, d_1\rangle, \langle u_2, d_2\rangle) = C(\langle u_1, d_2\rangle, \langle u_2, d_1\rangle)$.

The assumption means that the probability of including a sum of revenues inequality when it is valid for the assignment A_1 must be equal to the probability of including the reverse inequality when it is valid for the assignment $A_2 = (A_1 \backslash \{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\}) \cup \{\langle u_1, d_2\rangle, \langle u_2, d_1\rangle\}$. The probability C of choosing either inequality can be a function of the realizations of the firm characteristics in (u_1, u_2, d_1, d_2), but the probability must be the same whether the observed matches are $\{\langle u_1, d_1\rangle, \langle u_2, d_2\rangle\}$ or $\{\langle u_1, d_2\rangle, \langle u_2, d_1\rangle\}$.[10] Because all inequalities needed for identification are included in the limit as $H \to \infty$, sampling inequalities does not change point identification to set identification. In the empirical work, I sample each valid inequality with uniform probability within a market, which satisfies the relatively weak Assumption 4.3.

Often a researcher will not have a good idea of the boundaries in space and time of a matching market. By defining a market conservatively, so that the market definition used in estimation is weakly smaller than the true market, consistency will be maintained if the discarded inequalities are not necessary for point identification. Of course, throwing away valid inequalities might make the estimator less precise in a finite sample.

4.4 Consistency and Inference as the Number of Markets Grows

I first argue that the estimator that adds observations as the number of independent matching markets grows is consistent.

Assumption.

1. *The structural revenue function parameters* β *lie in a compact set* $\mathcal{B} \subseteq \mathbb{R}^{|\beta|}$, $|\beta| < \infty$.

[9] The estimator as $H \to \infty$ will not have a normal distribution. Therefore, I will avoid discussing how the choice of inequalities relates to statistical efficiency.

[10] Condition on the event that one of the two assignments A_1 and A_2 occurs. A weaker assumption is that, conditional on this event, the probabilities of the forward and reverse-direction inequalities must be the same.

2. *The elements of all X_{u_1,u_2,d_1,d_2} are not linearly dependent.*

3. *There is one element x_1 of X_{u_1,u_2,d_1,d_2} that has continuous support (induced by ρ) on the real line conditional on the other elements of X_{u_1,u_2,d_1,d_2}.*

4. *$\mathcal{B}$ is such that the coefficient β_1 on x_1 is normalized to be ± 1.*

5. *The assignment A is independently and identically distributed across markets.*

The assumption ruling out linear dependence relates to the informal discussion of identification above. The scale normalization that the coefficient of one regressor is ± 1 is innocuous because dividing by a positive scalar preserves an inequality. To operationalize the normalization, one maximizes the maximum score objective function imposing $\beta_1 = +1$ and then maximizes the objective function imposing $\beta_1 = -1$. The final set of estimates corresponds to the higher of the two objective function values. Some other assumptions are mentioned below.

Proposition 4.1. *Under the above assumption and the rank order property for many matching markets, as the number of markets $H \to \infty$, any $\hat{\beta}_H \in \mathcal{B}$ that maximizes the matching maximum score objective function* (9) *is a consistent estimator of $\beta^0 \in \mathcal{B}$, the parameter vector in the data generating process.*

The proof is the appendix. The proof uses the general consistency theorem for extremum estimators in Newey and McFadden (1994), which generalizes the early work of Manski (1975, 1985) on maximum score. The insight here is not the consistency proof, but the general idea that maximum score can be interpreted as a necessary-conditions approach for inequalities, at least for matching games with transferable utilities. Letting $\mathcal{A}$ be a set of A's, the maximum score estimator is consistent in part because of a law of large numbers, as

$$\operatorname{plim}_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} 1\left[A_h \in \mathcal{A}\right] = \int_{\mathcal{A}} \rho\left(A\right) dA = \Pr\left(\mathcal{A}\right),$$

where $1\left[A_h \in \mathcal{A}\right]$ equals 1 if an assignment in $\mathcal{A}$ occurs in market h.

The maximum score consistency proof shows that the true parameter vector β^0 maximizes the probability limit of the objective function. Such an argument would not work if the objective function involved minimizing the number of incorrect predictions times a penalty term (other than the current 1s and 0s) reflecting the difference $X_{u_1,u_2,d_1,d_2}{}'\beta$ between the left and right sides of the sum of revenues inequality, when evaluated at a hypothetical β. The rank order property suggests maximizing the number of correct inequalities, not allowing a violation in one inequality in order to minimize the degree of violation in another inequality.

Kim and Pollard (1990) show that the binary choice maximum score estimator converges at the rate of $\sqrt[3]{H}$ (instead of the more typical $\sqrt{H}$) and that its limiting distribution is too complex for use in inference. Abrevaya and Huang (2005) show that the bootstrap is inconsistent while Delgado, Rodríguez-Poo and Wolf (2001) show that another resampling procedure, subsampling, is consistent. Subsampling was developed by Politis and Romano (1994). The book Politis, Romano and Wolf

(1999) provides a detailed overview of subsampling. The empirical work on automotive suppliers uses subsampling for inference.

There are other options available to researchers. An alternative to subsampling is smoothing the indicator functions in the maximum score objective function. For the binary choice maximum score estimator, Horowitz (1992) proves that a smoothed estimator converges at a rate close to $\sqrt{H}$ and is asymptotically normal with a variance-covariance matrix than can be estimated and used for inference. Further, Horowitz (2002) shows the bootstrap is consistent for his smoothed maximum score estimator. Jun, Pinkse and Wan (2009) present a Chernozhukov and Hong (2003) Laplace type estimator (LTE). The LTE can converge at a rate close to $\sqrt{H}$; inference does not require a resampling procedure such as subsampling.

One can use set inference procedures for maximum score, even if the model is perhaps point identified. Point identification in maximum score is not equivalent to identification at infinity (Andrews and Schafgans, 1998). Rather, point identification involves finding firm characteristics such that $X_{u_1,u_2,d_1,d_2}{}'\beta^0 > 0 > X_{u_1,u_2,d_1,d_2}{}'\beta^1$, or the reverse, for the true parameter vector β^0 and some alternative $\beta^1 \neq \beta^0$. As β^0 is not known to the researcher, the full support condition on one element of X_{u_1,u_2,d_1,d_2} ensures that any needed values of X_{u_1,u_2,d_1,d_2} will be in the support of the data. A failure of this assumption results in set rather than point identification. Set identification is robust to the failure of support conditions for point identification. In a sense, set inference makes more use of the data. Bajari, Fox and Ryan (2008) explore set inference in maximum score, motivated by an industrial organization demand application. The set-identified subsampling approaches of Chernozhukov, Hong and Tamer (2007) and Romano and Shaikh (2010) can be used. The matching estimation software available on my website conducts subsampling inference for all of point- and set-identified maximum score and point- and set-identified maximum rank correlation (Santiago and Fox, 2009).

4.5 Consistency and Inference as the Number of Firms With Recorded Data In One Market Grows

I now turn to the case of $H = 1$, or estimation using one, typically large matching market. In this case, let J be the number of upstream firms with recorded physical partner lists in the assignment. As discussed earlier, the asymptotic argument here models the recorded observations on J upstream firms as a random sample from a true matching game with a continuum of firms. The objective function (9) can be rewritten, with a different normalizing constant, as

$$Q_J(\beta) = \frac{2}{J(J-1)} \sum_{u_1=1}^{J-1} \sum_{u_2=u_1+1}^{J} \sum_{\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\}\in I_{u_1,u_2}} 1\left[X_{u_1,u_2,d_1,d_2}{}'\beta \geq 0\right], \tag{10}$$

where I_{u_1,u_2} is the set of inequalities to include for the pair of upstream firms u_1 and u_2. For clarity, I have duplicated notation to use u_1 as both an index and as the characteristics of the corresponding upstream firm. The following assumption replaces the analogous assumption for many independent matching markets.

Assumption.

1. *The structural revenue function parameters β lie in a compact set $\mathcal{B} \subseteq \mathbb{R}^{|\beta|}$, $|\beta| < \infty$.*

2. *Each vector X_{u_1,u_2,d_1,d_2} corresponds to a pair of two physical partner lists and one downstream firm each from those lists. Each partner list p is i.i.d. with the density $g^{r^{\mathrm{up}},r^{\mathrm{down}},S}(p)$, where the functions are evaluated at their true values.*

3. *The elements of all X_{u_1,u_2,d_1,d_2} are not linearly dependent.*

4. *The vector X_{u_1,u_2,d_1,d_2} has at least one element x_1 with continuous support (induced by g) on the real line conditional on the other aspects of X_{u_1,u_2,d_1,d_2}.*

5. *$\mathcal{B}$ is such that the coefficient β_1 on x_1 is normalized to be ± 1.*

Proposition 4.2. *Under the above assumption and the rank order property for one large matching market, as the number of upstream firms with recorded data $J \to \infty$, any $\hat{\beta}_J \in \mathcal{B}$ that maximizes the maximum rank correlation objective function* (10) *is a consistent estimator of $\beta^0 \in \mathcal{B}$, the parameter vector in the data generating process.*

The proof in the appendix is largely omitted because of its similarity to the previous consistency proof. The estimator based on (10) is commonly called a maximum rank correlation estimator. Han (1987) introduced the estimator and showed consistency. Sherman (1993) shows that the maximum rank correlation estimator is $\sqrt{H}$-consistent and asymptotically normal. The objective function (10) at a given β is a U-statistic of second order. As H grows, the terms in the double summation grow proportionately to H^2. Intuitively, the inner summation acts like a smoother without requiring an explicit kernel and bandwidth. The derivation relies on a general set of results for the asymptotic distribution of U-processes in Sherman (1994).

The asymptotic distribution in Sherman (1993) requires potentially high-dimensional, nonparametric estimates of components of the variance matrix to be used in estimation. Subbotin (2007) proves that a resampling procedure, the bootstrap, is consistent for the maximum rank correlation estimator. It is interesting that the same objective function has two different asymptotic arguments for consistency. The two arguments lead to quite different statistical properties.

4.6 Comparison to Simulation Estimators

If one is willing to assume one of the two rank order properties holds, one has access to a consistent maximum score estimator that is computationally simple. The focal alternative to the maximum score estimator is likely some form of a simulation estimator, where the equilibria to matching games are computed as part of a nested procedure inside the evaluation of the statistical objective function. The most common simulation estimators are the method of simulated moments and maximum simulated likelihood. Both methods would be infeasible for the empirical work on automotive suppliers. The car parts data do not list the quotas of each firm, the maximum number of physical matches. Likewise, data on potential car parts suppliers or assemblers who in equilibrium are not matched are not available. Strong assumptions on these missing data would be needed to check whether an assignment could be part of a pairwise stable outcome.

If data on multiple matching markets are used, multiple pairwise stable equilibrium assignments become a serious issue. For simulation estimators, the only strategies to deal with multiple equilibria are extensions of work by Bajari, Hong and Ryan (2010) and Ciliberto and Tamer (2009). Both procedures require computer software to compute all equilibria to a game. In many-to-many matching games, there is no simple algorithm for computing either one or all pairwise stable outcomes or assignments in many-to-many matching, other than the often infeasible algorithm of checking every physically feasible assignment, one by one.[11] Even for one-to-one matching with 100 upstream and 100 downstream firms, there are many more assignments than the atoms in the universe. The computational cost of simulation estimators is exhibited in the Monte Carlo experiments in the next section.

Simulation estimators do have the advantage that explicit forms of unobserved heterogeneity can be included and the parameters of the distributions of heterogeneity can be estimated. This allows simulating probabilities of different equilibria, rather than just computing equilibria for particular values of unobservables.

There is a related literature on matching games without transferable utility; i.e. money is not used. Boyd, Lankford, Loeb and Wyckoff (2003), Sørensen (2007), and Gordon and Knight (2009) estimate Gale and Shapley (1962) matching games.[12] Multiple equilibria are typically even more numerous in non-transferable utility matching games although the above papers impose assumptions to work around multiplicity. The above papers use simulation estimators and are limited in the size of the matching markets they can consider.

5 Monte Carlo Experiments

This section presents evidence that the maximum score estimator works well in finite samples and with i.i.d., non-logit, match-specific errors. This section reports a Monte Carlo study for an estimator that has not been proved to be formally consistent: the rank order property does not hold. The Monte Carlo study examines games of one-to-one, two-sided matching with finite numbers of agents in the true model. I choose the simple case of one-to-one matching for computational reasons: to make it easier to generate the fake data and especially to make it easier to compare the maximum score estimator to alternative likelihood and method of moments simulation estimators. As I have discussed, one-to-one matching is a sufficient but not necessary condition to rule out multiple equilibrium assignments.

5.1 Varying Sample Size and Error Dispersion

Each agent is distinguished by two characteristics, for upstream firm u, u^1 and u^2, and for downstream firm d, d^1 and d^2. The distribution of each $u = (u^1, u^2)$ and each $d = (d^1, d^2)$ is bivariate normal, with means of 1, standard deviations of 1, and covariances between u^1 and u^2 and between d^1 and d^2

[11] Checking whether an assignment may be part of a pairwise stable outcome requires searching for a corresponding set of transfers that satisfy the definition of pairwise stability. Checking whether the sum of revenues inequalities are satisfied is not enough.

[12] Hitsch, Hortaçsu and Ariely (2009) use data on both desired and rejected matches to estimate preferences without using an equilibrium model. They then find that a calibrated model's prediction fits observed matching behavior. Echenique (2008) examines testable restrictions on the lattice of equilibrium assignments of the Gale and Shapley (1962) model.

of 1/2. The nonzero covariance suggests a multivariate estimator might give different estimates than a univariate estimator. In one-to-one matching, it is difficult to distinguish the functions r^{up} and r^{down}, as what matters for pairwise stability, absent the individual rationality decision to be unmatched, is the total production $f(\langle u,d\rangle) = r^{\text{up}}(\langle u,d\rangle) + r^{\text{down}}(\langle u,d\rangle)$ from each match. Therefore, I primitively specify the production to each match as

$$f_{\beta_1,\beta_2}(\langle u,d\rangle) + \epsilon_{\langle u,d\rangle} = \beta_1 u^1 d^1 + \beta_2 u^2 d^2 + \epsilon_{\langle u,d\rangle},$$

where $\epsilon_{\langle u,d\rangle}$ is a match-specific unobservable with a distribution varied in the experiments. The true parameter values are $\beta_1 = 1.0$ and $\beta_2 = 1.5$, so that the second observable characteristic is more important in sorting. The sign of β_1 is superconsistently estimable, so in maximum score I set it to the true value of +1.[13] The parameter value, not just the sign, of β_1 is estimated in the parametric likelihood and method of moments simulation estimators. In those estimators, the scale normalization is on the standard deviation of $\epsilon_{\langle u,d\rangle}$ and not on a parameter.

To generate finite data, I sample match specific errors and solve for the optimal assignment using a linear programming problem described in Roth and Sotomayor (1990). The linear programming formulation ensures that all consummated matches provide non-negative surplus.

Table 1 demonstrates that the bias and root mean-squared error (RMSE) of the matching maximum score estimator decrease with sample sizes in the experiments considered. There are two notions of sample size: the number of upstream firms in a single market (equal to the number of downstream firms for simplicity) and the number of markets. The true distribution of $\epsilon_{\langle u,d\rangle}$ is a mixture of two normal distributions, given in the footnote to the table. The choice of a bimodal distribution highlights the nonparametric treatment of the error distribution in maximum score estimation. The right panel of Table 1 uses a standard deviation for $\epsilon_{\langle u,d\rangle}$ that is ten times higher than the left panel's standard deviation. In the right panel, the distributions of $u = (u^1, u^2)$ and of $d = (d^1, d^2)$ are such that most explanatory power for the total production of a match comes from the error term. The $\epsilon_{\langle u,d\rangle}$ term has a standard deviation of 10 while the explanatory portion of the model, $\beta_1 u^1 d^1 + \beta_2 u^2 d^2$, has a standard deviation of 3.68 at the true parameters. The $\epsilon_{\langle u,d\rangle}$ term will have a standard deviation up to 50 in Table 2.

In the first row of the left panel of Table 1, the bias and RMSE are relatively high for 3 downstream and 3 upstream firms (6 total) for each market and 100 markets. The bias of -0.12 is manageable compared to a true value of $\beta_2 = 1.5$, as is the RMSE of 0.66. The bias and RMSE are slightly smaller for 10 firms on each side of the market and only 10 markets. Both the bias and RMSE decrease when more firms are added to each market: the third row reports 30 firms on each side and 10 markets. The bias remains about the same while the RMSE decreases further with 60 firms on each side and 10 markets. The fifth row then shows that increasing the number of markets to 40 almost eliminates the bias and further reduces the RMSE.

Another question is how well the estimator works in a finite sample with data on only one fairly large matching market. The sixth row of the left panel uses 100 firms on each side of the market, but

[13] For each replication for maximum score, the Monte Carlo study reports the maximizer $\hat{\beta}_2$ provided by the optimization routine. If the maximum reported by the optimization package tends to always be near the lower bound of the set of finite-sample maxima, it could create an apparent downward, finite-sample bias. In practice, the range of global maxima is small.

only one market. The bias and RMSE are relatively low. The bias and RMSE then decline in the seventh row as the number of firms on each side increases to 200.

As expected, the bias and RMSE are larger in the right panel of Table 1, when the standard deviation of the additive error $\epsilon_{\langle u,d\rangle}$ is increased by a factor of 10. There is less signal in the data when unobservables drive matching. However, as before, the RMSE and the bias go down with both measures of sample size. The bias in particular is relatively small with higher sample sizes. In these experiments, the estimator is not very biased when there are i.i.d. match-specific errors. This supports the use of the maximum score estimator even when it may be formally misspecified, as when there are i.i.d. match-specific errors and the truth is not the Choo and Siow (2006) logit matching model. The misspecification is analogous to estimating a single-agent logit when the true model is probit much more than not correcting for selection bias or omitted variable bias. This misspecification bias is relatively small in the considered experiments.

5.2 Comparing Maximum Score to Parametric Estimators

Table 2 compares maximum score to a likelihood and to a method of moments estimator. Both the likelihood and method of moments estimators are parametric in that they impose a known distribution for $\epsilon_{\langle u,d\rangle}$: the distribution is assumed to be normal in estimation. The top panel in Table 2 lets the true distribution indeed be normal, with increasing levels of dispersion. The bottom panel of Table 2 considers the case where the true distribution is a mixture of two normals, so that the parametric estimators are misspecified and hence inconsistent. Maximum score itself is often misspecified when the model has i.i.d. errors at the match level and the true model is not Choo and Siow (2006).

The implementation details of the two parametric estimators are many and available from the author upon request, but both the likelihood and method of moments estimators involve simulation. Some effort was put into tuning each of the parametric estimators. A straightforward simulated likelihood estimator that was first implemented suffered from a serious tradeoff between insurmountable simulation errors and computational costs. Therefore, I turned to a frequentist, data augmentation MCMC implementation of maximum likelihood, following Jacquier, Johannes and Polson (2007). The data augmentation scheme draws the latent production values for each match to be consistent with the observed assignment in the data, which dramatically improves the performance of maximum likelihood. The method of moments fits the sample covariances of the form $\text{Cov}\left(u^1, d^2\right)$ (four moments in total), as seen in the matches in the data and in the R computed equilibrium assignments (using linear programming) for each market. The scalar R is the number of sets of simulation draws (there is one error for each potential match). The method of simulated moments is consistent as the number of markets increases for a fixed R. Table 2 uses five sets of draws for each market, which means the equilibrium assignment is computed five times for each market in order to evaluate the objective function.

The first main results in Table 2 are the run times of each of the estimators. The first row of the table considers a dataset of 100 independent matching markets with 3 upstream and 3 downstream firms in each market. This is a trivial problem for maximum score, taking 2 seconds on average to estimate. The MCMC likelihood estimator took 2700 seconds, which actually is a lot less than a straightforward simulated likelihood estimator without data augmentation and with low simulation error would take. The speed could be increased by using fewer MCMC iterations, but the robust finding is that 2700 seconds is several orders of magnitude slower than 2 seconds. Likewise, the five sets of simulation draws for each market in the method of moments lead to a run time of 1400 seconds. Fewer than five sets of draws would increase speed at the expense of statistical performance, but again the robust finding is that 1400 seconds is several orders of magnitude slower than 2 seconds. When the number of firms in each side of each market is increased from 3 to 4, maximum score still takes 2 seconds on average, while the likelihood procedure takes 7100 seconds and the method of moments estimator takes 2100 seconds.

Table 2 also looks at the statistical performance of the two parametric estimators and maximum score. In the upper panel, with the true data being generated by normal errors, the parametric estimators are consistent and the maximum score estimator is misspecified. The rows refer to different sample sizes and dispersions of the error terms. As expected, maximum score usually has a higher bias in absolute value. When the normal errors have a small dispersion of 1, the parametric estimators also have lower RMSEs. When the normal standard deviation increases to 25, the maximum score estimator has a lower RMSE than the method of moments estimator. When the normal standard deviation increases to 50, the maximum score estimator has a lower RMSE than both the likelihood and method of moments estimators. In particular, the simulated GMM estimator has high RMSE. In these experiments, the semiparametric maximum score estimator performs relatively well statistically (low RMSE) when the signal in the data is quite low relative to the noise (the magnitude of the error terms).

The lower panel of Table 2 considers experiments where the errors have a mixture of normals distribution, so that the parametric estimators are misspecified and inconsistent. Although not universal, the RMSEs in the lower panel tend to be higher than for the equivalent cases in the upper panel. The absolute value of the biases are less reliably higher in the second panel. In some experiments, the semiparametric maximum score estimator has a lower bias or RMSE than the also misspecified parametric estimators. The two parametric estimators are particularly biased when the standard deviation of the error terms is small. Perhaps the misspecified functional form for the distribution of the errors plays a greater role in the point estimates when the standard deviation of the error terms is small.

Table 2 considers only experiments with 3 or 4 upstream and 3 or 4 downstream firms in each of 100 markets. These are trivially small matching markets compared to those of interest to many researchers in industrial organization. The introduction discusses how one component category in the automobile market has 2627 different car parts, the equivalent of a firm in matching theory. There is no hope that a parametric estimator could be computationally implemented for a matching market with such a large number of firms. The parametric estimators suffer from a curse of dimensionality that arises from having to solve a matching game repeatedly (simulated method of moments) or to draw error terms from increasingly complicated inequalities (data augmentation MCMC likelihood). This is documented in Figure 1, which plots the number of firms on the horizontal axis and the run time of the method of moments procedure on the vertical axis. The relationship is indeed convex, so, as expected, simulated GMM does suffer from a computational curse of dimensionality. Only numbers of upstream firms up to 11 are considered for computational reasons.

6 Data on Matching in the Car Parts Industry

I now present an empirical application about the matching of suppliers to assemblers in the automobile industry. Automobile assemblers are well-known, large manufacturers, such as BMW, Ford or Honda. Automotive suppliers are less well-known to the public, and range from large companies such as Bosch to smaller firms that specialize in one type of car part. A car is one of the most complicated manufacturing goods sold to individual consumers. Making a car be both high quality and inexpensive is a technical challenge. Developing the supply chain is an important part of that challenge. More so than in many other manufacturing industries, suppliers in the automobile industry receive a large amount of coverage in the industry press because of their economic importance.

A matching opportunity in the automotive industry is an individual car part that is needed for a car model. A particular part l in the data is attached to an assembler, d. Therefore a physical match in this industry is a triple $\langle u, d, l\rangle$. The same supplier can supply more than one part to the same assembler: $\langle u, d, l_1\rangle$ and $\langle u, d, l_2\rangle$ represent two different matches (car parts) between assembler d and supplier u. This is a two-sided, many-to-many matching game between assemblers and suppliers, with the added wrinkle that a supplier can be matched to the same assembler multiple times.

The data come from SupplierBusiness, an analyst firm. There are 1252 suppliers, 14 parent companies, 52 car brands, 392 car models, and 52,492 car parts. While the data cover different model years, for simplicity I ignore the time dimension and treat each market as clearing simultaneously.[14] The data group car parts into component categories, and I treat each component category as a statistically independent matching market.[15] I only use component categories for which there are more than 100 possible inequalities. Eliminating the small categories results in 187 distinct component categories, such as pedal assembly and coolant/water hoses. I assume any nonlinearities between multiple matches involving the same supplier occur only within component categories; there are no spillovers across the different matching markets. A triplet $\langle u, d, l\rangle$ in the data then could be the front pads of a Fiat 500 (a car) supplied by Federal-Mogul. Front pads are in the component category (matching market) disk brakes.[16]

One of the empirical applications focuses on General Motors divesting Opel, a brand it owns in Europe. In order to model the interdependence of the European and North American operations of General Motors and suppliers to General Motors, the definition of a matching market is car parts in a particular component category used in cars assembled in Europe and North America. Most of the assemblers and many of the larger suppliers operate on multiple continents.[17] However, the point estimates found when splitting Europe and North America into separate matching markets are similar to those presented here, suggesting that geographic market definitions do not play a large role

[14] Car models are refreshed around once every five years.

[15] The same firm may appear in multiple component categories, and so a researcher might want to model spillovers and hence statistical dependence in the outcomes across component categories. Pooling component categories poses no issue with the econometric method. The history of the industry shows that many US suppliers were formed in the 1910's and 1920's around Detroit (Klier and Rubenstein, 2008). Some firms chose to specialize in one or a few component categories and others specialized in more component categories. The particular historical pattern of what component categories each supplier produces lies outside of the scope of this investigation.

[16] The parameter estimates in this paper would presumably change if SupplierBusiness aggregated or disaggregated car parts in different ways.

[17] Nissan and Renault are treated as one assembler because of their deep integration. Chrysler and Daimler were part of the same assembler during the period of the data.

in identifying the parameters. Note that many of the estimated gains to specialization to a supplier likely come from plant co-location: using one supplier plant to supply the same type of car part to multiple car models assembled in the same plant or in nearby plants. Thus, an empirical regularity of certain suppliers being more prevalent in one continent than another is consistent with the gains to specialization that I seek to estimate.[18]

The data have poor coverage for car models assembled in Asia, so I cannot include the corresponding car parts in the empirical work. I do focus heavily on car parts used on cars assembled in Europe and North America by assemblers with headquarters in Asia.

The automotive supplier empirical application is a good showcase for the strengths of the matching estimator. The matching markets modeled here contain many more agents than the markets modeled in most other papers on estimating matching games. The computational simplicity of maximum score, or some other approach that avoids repeated computations of model outcomes, is needed here. Other than my related use of the estimator in Fox and Bajari (2010), this is the first empirical application to a many-to-many matching market where the payoffs to a set of matches are not additively separable across the individual matches. I focus on specialization in the portfolio of matches for suppliers and assemblers. Finally, matched firms exchange money, but the prices of the car parts are not in publicly available data. The matching estimator does not require data on the transfers, even though they are present in the economic model being estimated. Likewise, data on potential entrants to each component category and to automobile assembly are not needed.

7 The Costs of Assemblers Divesting Brands

7.1 General Motors and Opel

In 2009, General Motors (GM), the world's largest automobile assembler for most of the twentieth century, declared bankruptcy. As part of the bankruptcy process, GM divested or eliminated several of its brands, including Pontiac and Saturn in North America and SAAB in Europe. Economists know little about the benefits and costs of large assemblers in the globally integrated automobile industry divesting brands. This paper seeks to use the matching patterns in the car parts industry to estimate one aspect of the costs of divestment.

A major public policy issue during 2009 was whether General Motors should also divest its largest European brands, Opel and Vauxhall.[19] Opel is based in Germany and Vauxhall is based in the United Kingdom. Over the period of the data, Opel also had assembly plants in Belgium, Hungary, Poland, and Russia. Consistent with the close link between Opel and Vauxhall, they will be grouped together as one brand, Opel, in the empirical work.

A major advocate of GM divesting Opel was the German government, which desired to protect jobs at Opel assembly plants, at Opel dealers and at suppliers to Opel, but was reluctant to subsidize

[18] A few suppliers are owned by assemblers. I ignore this vertical-integration decision in my analysis, in part because I lack data on supplier ownership and in part because vertical integration is just an extreme version of specialization, the focus of my investigation. If a supplier sends car parts to only one assembler, that data are recorded and used as endogenous matching outcomes. Vertical integration in automobile manufacturing has been studied previously (Monteverde and Teece, 1982; Novak and Eppinger, 2001; Novak and Stern, 2008, 2009).

[19] GM has owned Opel since 1929, although its control temporarily lapsed during the second World War.

a bankrupt North American firm. During most of 2009, the presumption by GM was that Opel would be divested. Indeed, GM held an auction and and agreed to sell Opel to a consortium from Canada and Russia. In November 2009, GM canceled the sale and kept Opel as an integrated subsidiary of GM. Opel and the North American operations of GM share many common platforms for basing individual models on. One reason for keeping Opel integrated is that a larger, global assembler will have gains from specialization in its own assembly plants and in the plants of suppliers. Increasing the gains to suppliers from specializing in producing car parts for GM may indirectly benefit GM through lower prices for car parts.

7.2 Structural Revenue Functions

7.2.1 Revenue Functions for Suppliers

This section estimates the structural revenue functions of assemblers and of suppliers for the portfolio of car parts each firm sources or supplies. The revenue functions take as arguments measures of how specialized each portfolio of car parts is at several levels. For suppliers, the revenue function specification says suppliers may specialize in four areas: parts (in the same component category) for an individual car, parts for cars from a particular brand (Chevrolet, Audi), parts for cars from a particular parent company or assembler (General Motors, Volkswagen) and parts for cars for brands with headquarters on a particular continent (Asia, Europe, North America).

The pattern of sorting in the car parts market is used to measure the relative importance of specializing at different levels of aggregation. The management literature has suggested that supplier specialization may be a key driver of assembler performance (Dyer, 1996, 1997; Novak and Wernerfelt, 2007).

When I consider the counterfactual of GM divesting Opel and making it an independent assembler or parent company, the changes in total structural revenue will be generated by the estimated parameter on the importance of specialization at the parent company level, relative to the values of the other parameters.

Fix a component category or matching market. Let M be a portfolio of car part matches for supplier u, where an element of M is $\langle u, d, l \rangle$. The structural revenue function of suppliers is

$$r^{\text{up}}_{\beta^{\text{up}}}(M) = \beta^{\text{up}}_{\text{Cont.}} z^{\text{up}}_{\text{Continent}}(M) + \beta^{\text{up}}_{\text{PG}} z^{\text{up}}_{\text{ParentGroup}}(M) + \beta^{\text{up}}_{\text{Brand}} z^{\text{up}}_{\text{Brand}}(M) + \beta^{\text{up}}_{\text{Model}} z^{\text{up}}_{\text{Model}}(M),$$

where $\beta^{\text{up}} = (\beta^{\text{up}}_{\text{Cont.}}, \beta^{\text{up}}_{\text{PG}}, \beta^{\text{up}}_{\text{Brand}}, \beta^{\text{up}}_{\text{Model}})$ is a vector of estimable parameters and where

$$Z^{\text{up}}(M) = \left(z^{\text{up}}_{\text{Continent}}(M), z^{\text{up}}_{\text{ParentGroup}}(M), z^{\text{up}}_{\text{Brand}}(M), z^{\text{up}}_{\text{Model}}(M)\right)$$

is a vector of observable characteristics of the portfolio M of car part matches. Each component of $Z^{\text{up}}(M)$ is a measure of specialization in one level of opportunities to sell car parts. The choice of a measure of specialization is somewhat arbitrary. I use the Herfindahl-Hirschman Index (HHI) because economists are familiar with its units, which range between 0 and 1. For example, say the North

American firms of Chrysler, General Motors and Ford are the only three assemblers. Then

$$z^{\text{up}}_{\text{ParentGroup}}(M) = \left(\frac{\#\,\text{Chrysler parts in } M}{\#\,\text{total parts in } M}\right)^2 + \left(\frac{\#\,\text{Ford parts in } M}{\#\,\text{total parts in } M}\right)^2 + \left(\frac{\#\,\text{GM parts in } M}{\#\,\text{total parts in } M}\right)^2. \tag{11}$$

As this specialization measure enters the revenue function for a supplier, $z^{\text{up}}_{\text{ParentGroup}}(M)$ is 1 if the supplier sells parts only to, say, GM and 1/3 if it sells an equal number of parts to each assembler. The use of the HHI is different than in antitrust; here the HHI is a measure of specialization for a portfolio M of car parts for a particular supplier u and is not a measure of concentration in the overall market for car parts. The specialization measure $z^{\text{up}}_{\text{ParentGroup}}(M)$ can be computed both for the matches M_u for supplier u in the data and in the counterfactual matches in the sum of revenues inequality (8).

Let (d, l) be a particular car part matching opportunity, for part l on a car model sold by assembler d. In the notation used earlier, the characteristics (d, l) represents are the identity of the car model, the brand of the car model, the parent company of the brand, and the continent of headquarters of the brand.[20] The vector $Z^{\text{up}}(M)$ is a nonlinear transformation of the underlying characteristics of car part matching opportunities. By construction, two parts for the same car also have the same brand, parent group and continent. Two car parts for cars from the same brand are automatically in the same parent group and the brand only has one headquarters, so the parts are from a brand with a headquarters in the same continent as well. Two cars from the same parent group are not necessarily from the same continent, as Opel is a European brand of GM and Chevrolet is a North American brand of GM.

The four specialization measures in $Z^{\text{up}}(M)$ are highly correlated. Just as univariate linear least squares applied to each covariate separately produces different slope coefficients than multivariate linear least squares when the covariates are correlated, a univariate matching theoretic analysis (such as Becker (1973)) on each characteristic separately will be inadequate here. A univariate analysis of say $\beta^{\text{up}}_{\text{PG}} z^{\text{up}}_{\text{ParentGroup}}(M)$ would just amount to saying that $\beta^{\text{up}}_{\text{PG}} > 0$ when each supplier does more business with certain parent groups than others. In principle, even this conclusion about the sign of $\beta^{\text{up}}_{\text{PG}}$ could be wrong if the correlation with the other three characteristics is not considered in estimation. Here I measure the relative importance of each of the four types of specialization: at which level do the returns to specialization occur? This requires formal statistical analysis to estimate the vector β^{up}.

7.2.2 Revenue Functions for Assemblers

The structural revenue function of assemblers has a similar functional form, focusing on specializing in a small number of suppliers. Let M now be a portfolio of car parts, where all car parts have the assembler or parent group, d. The revenue function is $r^{\text{down}}_{\beta^{\text{down}}}(M) = Z^{\text{down}}(M)' \beta^{\text{down}}$, where

[20] Many other upstream firm characteristics would be endogenous at the level of the equilibrium matching considered here. For example, many of the benefits of specialization occur through plant co-location and so suppliers and assembler plant locations should be considered endogenous matching outcomes rather than exogenous firm characteristics. With just-in-time production at many assembly sites, supplier factories are built short distances away so parts can be produced and shipped to the assembly site within hours, in many cases.

$\beta^{\text{down}} = \left(\beta^{\text{down}}_{\text{PG}}, \beta^{\text{down}}_{\text{Brand}}, \beta^{\text{down}}_{\text{Model}}\right)$ is a vector of estimable parameters and where

$$Z^{\text{down}}(M) = \left(z^{\text{down}}_{\text{ParentGroup}}(M), z^{\text{down}}_{\text{Brand}}(M), z^{\text{down}}_{\text{Model}}(M)\right)$$

is a vector of specialization measures for the assembler. For conciseness, I do not include a term for specialization at the continent of brand headquarters level.

The term $z^{\text{down}}_{\text{ParentGroup}}(M)$ is a Herfindahl index for the concentration of suppliers selling parts to the assembler d. Given a portfolio of car parts M, let $\iota(M)$ be the set of suppliers u who sell at least one car part in M to d. Then

$$z^{\text{down}}_{\text{ParentGroup}}(M) = \sum_{u \in \iota(M)} \left(\frac{\#\,\text{parts sold by supplier}\,u\,\text{in}\,M}{\#\,\text{total parts in}\,M}\right)^2.$$

Next, $z^{\text{down}}_{\text{Brand}}(M)$ is the mean of such a Herfindahl index computed for each brand separately. Say d is GM and the only two brands of GM are Chevrolet (Chevy) and Opel and let $\iota(M, \text{Opel})$ be the set of suppliers selling parts to Opel in M. Then, for GM,

$$\begin{aligned} z^{\text{down}}_{\text{Brand}}(M) = \frac{1}{2} \sum_{u \in \iota(M,\text{Opel})} & \left(\frac{\#\,\text{parts sold by supplier}\,u\,\text{to Opel in}\,M}{\#\,\text{total parts for Opel in}\,M}\right)^2 \\ & + \frac{1}{2} \sum_{u \in \iota(M,\text{Chevy})} \left(\frac{\#\,\text{parts sold by supplier}\,u\,\text{to Chevy in}\,M}{\#\,\text{total parts for Chevy in}\,M}\right)^2. \end{aligned}$$

Likewise, $z^{\text{down}}_{\text{Model}}(M)$ is the mean across car models sold by GM of the Herfindahl index calculated for the suppliers of parts to each car model separately. As with suppliers, $z^{\text{down}}_{\text{Model}}(M)$ can be evaluated at the actual set of suppliers in the data, captured in M_d, and counterfactual portfolios in sum of revenues inequalities.

The underlying supplier characteristic (in each u) that enters $Z^{\text{down}}(M)$ is simply the identity of each supplier. The characteristics of car parts that enter $Z^{\text{down}}(M)$ are the number of parts, car model, brand and parent company. The set of car parts M_u for upstream firm u is different than the set of car parts M_d for downstream firm d. This and the nonlinear construction of $Z^{\text{down}}(M)$ and $Z^{\text{up}}(M)$ allow the separate identification of β^{up} and β^{down}, as discussed in section 4.1.

The sum of revenues inequalities used in estimation keep the number of car parts produced by each supplier (and, more obviously, the set of car parts needed on each car model) the same. With strong returns to specialization, it may be more efficient to have fewer but individually larger suppliers. The optimality of supplier size is not imposed as part of the estimator. Nor can the gains from assembler scale be identified from a sum of revenues inequality, if each car part and each car model are weighted equally. This paper models the car parts market, not the market for corporate control of car brands and car models. Not imposing the optimality of supplier and assembler sizes might be an advantage, as other concerns such as capacity constraints and antitrust rules could limit firm size. On the other hand, one of the benefits of GM not divesting Opel is keeping a larger scale, and the sum of revenues inequalities do not identify a pure scale economy for GM owning Opel. Instead, I focus on the gains to assemblers and particularly to suppliers from specialization, for a fixed number of car parts.

7.3 Point Estimates for Revenue Functions

Table 3 presents the point estimates and confidence intervals for the structural revenue functions for upstream and downstream firms. I randomly sample a maximum of 2000 inequalities per component category. All theoretically valid inequalities with two different suppliers are sampled with an equal probability, which satisfies Assumption 4.3. Other details are in the footnote to Table 3.

The parameter $\beta_{\text{Cont.}}$ is normalized to be ± 1. The other parameters in Table 3 are interpreted relative to $\beta_{\text{Cont.}}$. The most important finding in that the point estimates of the assembler parameters in β^{down} have a much lower order of magnitude than the supplier parameters in β^{up}. This is not because of a difference in the units of Z^{up} and Z^{down}; the rightmost columns of Table 3 report the means and standard deviations of the specialization measures for both suppliers and assemblers. The specialization (HHI) measures are about the same magnitudes for both suppliers and assemblers. The assembler point estimates show that assemblers dislike sourcing their supplies from only a few suppliers. However, the magnitude of any such effect is quite small and the hypothesis that an assembler point estimate is 0 would only be rejected for assembler specialization (having a narrow supplier base) at the parent group level. What is possibly explaining the small magnitude effects is that two economic forces may offset each other: assemblers prefer to have a diverse supplier base to avoid placing their success in the hands of one supplier (hold up) while there may be some manufacturing benefits from having a fewer number of suppliers. Regardless, the point estimates show that assembler specialization is much less important than supplier specialization in explaining outcomes. One caveat is that the confidence intervals for assembler specialization at the brand and model levels do contain larger, in absolute value, coefficient magnitudes.

For suppliers, Table 3 shows that all four estimates in β^{up} are positive, meaning as expected specialization on these dimensions increases the revenue of suppliers. The estimated parameters show that a given level of specialization at the parent-group level is 4.1 times more important in revenue than the same level of specialization at the continent-of-brand-headquarters level. Most specialization benefits occur within firm boundaries rather than across them. At the same time, the standard deviation of parent-group-specialization HHI, from each supplier's viewpoint, is 0.29, meaning the variation in parent-group specialization across suppliers is high. A naive researcher might be inclined to interpret this dispersion as evidence parent-group specialization is unimportant. This would be wrong: the maximum score estimator accounts for the fact that more available matching opportunities occur across firm boundaries than within them. An estimate of a structural parameter such as the coefficient on parent group tells us the importance of parent group in the structural revenue from a set of supplier relationships.

Table 3 also shows that specialization at the brand and model levels is even more important than specialization at the parent-group level, although the brand and parent-group confidence intervals overlap. The high point estimate of 86 for model specialization likely comes from supplier and assembler plant co-location: car models of even the same brand may be built in separate plants and some benefits from specialization may occur from saving on the need to have multiple supplier plants for each model. Also, the technological compatibility of car parts occurs mainly at the model level. Notice how the standard deviation of the HHI-specialization measure is about the same (around 0.3) for the parent-group, brand and model measures, and how the mean HHI declines from parent group to brand to model. Again, naive researchers might use the means to conclude that specialization at the model level is less important or use the standard deviations to conclude that specialization at all three levels are equally important. The structural estimates of the revenue functions give statistically consistent estimates of the relative importance of the types of specialization in the structural revenue functions for supplier relationships.

Table 3 also shows that there are 332,218 inequalities used in estimation. Of those, 81% are satisfied at the reported point estimates. The fraction of satisfied inequalities is a measure of statistical fit, which appears to be good in this set of estimates.

7.4 Supplier Revenue Loss From GM Divesting Opel

Encouraging General Motors to divest Opel was a major policy issue in Germany during 2009. The revealed preference of GM to back away from selling Opel to outside investors suggests that GM felt that Opel was important to its performance. One possibility is that GM feared a loss of economies of scale (total size) or scope (strength in fuel efficient cars that could be transferred from Europe to North America, say) from such a divestiture. Matching in the car parts market is not necessarily informative about assembler economies of scale and scope.

Using information from the car parts market, and in particular in light of the minuscule point estimates on assembler specialization above, the major estimated effect of GM divesting Opel will come from suppliers to GM being less specialized as GM's and Opel's models technologically diverge. This will hurt GM through equilibrium transfers: suppliers will charge a higher price to GM. In each component category, I construct the counterfactual sum of structural revenue to suppliers if Opel and the rest of GM are now treated as separate assemblers, or parent groups. The same upstream firms supply the same car parts to the same car models, but now the Opel models are produced by an independent parent group. In (11), some parts are transferred to a new parent group and so the measure of parent group specialization weakly decreases for any supplier that sells any parts to Opel. The decrease in $\beta^{\text{up}}_{\text{PG}} z^{\text{up}}_{\text{ParentGroup}}(M)$ gives the decrease in the structural revenue for each supplier who sells at least one part to Opel. I focus on a percentage decrease measure

$$\frac{\beta^{\text{up}}_{\text{PG}} \Delta z^{\text{up}}_{\text{ParentGroup}}(M_u)}{r^{\text{up}}_{\beta^{\text{up}}}(M_u)},$$

for a particular upstream firm with the matches M_u in the data. Note that this measure imposes a cardinal (up to scale) interpretation of a supplier's revenue function, as opposed to identifying a supplier's revenue function only up to a positive monotonic transformation. Fox (2010) proves that the cardinal aspects of a related function are identified nonparametrically in matching games with transfers.

Table 4 reports statistics for the distribution of percentage changes in structural revenue for suppliers. A supplier in the table is a real-life supplier in a particular component category. Only suppliers who sell at least one part to Opel and one car part to another GM brand are affected and so included in the table. The mean loss is quite small, at 1%. This reflects suppliers where either Opel is a small fraction of car parts or a very large fraction of parts, so GM divesting Opel makes little difference in how specialized the supplier is. The overall loss is small; the 0.10 quantile of losses is 3% of structural revenue. This result follows from the parameter estimates in Table 3, where the point estimates for the coefficients on brand and especially model specialization are several times larger than the coefficient on parent group specialization. At the extreme, some suppliers lose a fair bit of revenue: the largest loss is 9%.[21]

8 The Benefits to Domestic Suppliers From Foreign Assemblers

European and North American countries have imposed formal and political-pressure based trade barriers to imports of automobiles from Asia. Consequently, most Asian assemblers who sell cars in Europe and North America assemble cars in Europe and North America as well. While some parts are imported from Asia, Asian assembly plants in Europe and North America use many parts produced locally as well (perhaps because of more political pressure). As Klier and Rubenstein (2008) document for Asian assemblers in North America, a key part of operating an assembly plant is developing a network of high-quality suppliers.

Despite some occasional quality setbacks, the magazine *Consumer Reports* and other sources routinely record that brands with headquarters in Asia (Japan, Korea) have higher quality automobiles than brands with headquarters in Europe or North America. The parts supplied to high-quality cars must also be of high quality. Liker and Wu (2000) document that suppliers to Japanese-owned brands in the US produce fewer parts requiring reworking or scrapping, for example. Because of this emphasis on quality, the suppliers to, say, Toyota undergo a rigorous screening and training program, the Supplier Development Program, before producing a large volume of car parts for Toyota (Langfield-Smith and Greenwood, 1998). Indeed, there is a hierarchy of suppliers, with more trusted Toyota suppliers being allowed to supply more car parts (Kamath and Liker, 1994; Liker and Wu, 2000).

It is possible that the need by Asian assemblers for higher-quality suppliers benefits the entire domestic supplier bases in Europe and North America. If a supplier is of high-enough quality to deal with an Asian assembler, non-Asian assemblers that also source parts from that supplier may also benefit. If this potential effect is causal (the suppliers were not of sufficiently high quality before the Asian assemblers' entry), it is evidence that trade barriers that promote Asian-owned assembly plants in Europe and North America may indirectly aid non-Asian (domestic) assemblers, as those producers now have access to higher-quality suppliers. This is an underexplored channel by which foreign-direct investment in assembly plants may raise the quality of producers in upstream markets. Indeed, there is evidence in the management literature that Asian assemblers do causally upgrade the quality of their suppliers: the Supplier Development Program mentioned above, for example (Langfield-Smith and Greenwood, 1998).

This section complements the management literature by providing evidence from sorting in the market for car parts that is consistent with suppliers to Asian assemblers being higher quality than

[21] I compute but do not report the small changes in GM's and Opel's revenues from divesting Opel. Because the coefficient estimates on assembler specialization in Table 3 are small in magnitude, the overwhelming effect is estimated to be on suppliers.

other suppliers. Measures of car part quality by individual suppliers are presumably observed by assemblers, but are not publicly available. In this section, a measure of quality will be a supplier's share of the market for supplying parts to Asian assemblers. If Asian assemblers together demand 100 parts in a particular component category, and one firm sells 30 of them, its quality measure will be 0.30. In notation,

$$u^1 = \frac{\#\,\text{Asian assembler parts supplied by } u}{\text{total}\,\#\,\text{Asian assembler parts all suppliers}},$$

for supplier $u = \left(u^1, \ldots\right)$. This is not a specialization measure, as a firm could sell many parts to Asian assemblers and many parts to non-Asian assemblers. This quality measure is treated as an exogenous characteristic of supplier u. If u^1 were included without interactions in a revenue function, it would difference out of the sum of revenues inequalities (4). Instead, the estimated supplier revenue function includes an interaction of u^1 with specialization by the continent headquarters of the brand, discussed earlier:

$$z^{\text{up}}_{\text{CompAdv}}(M) = u^1 \cdot z^{\text{up}}_{\text{Continent}}(M),$$

where all matches in M involve supplier u and the abbreviation is short for "competitive advantage". The interpretation of the corresponding supplier parameter $\beta^{\text{up}}_{\text{CompAdv}}$, if it is estimated to be negative, is that suppliers with higher u^1 (greater shares of the market for supplying Asian assemblers) gain less benefit from selling parts to only one type of assembler than firms with lower u^1. Thus, firms with higher Asian shares can go out and win business from non-Asian assemblers, which is consistent with those firms have a competitive edge (possibly from higher quality parts) over other suppliers. The empirical pattern in the data will be that suppliers with high u^1 have diverse (across continents of assembler origin) portfolios of car parts that they supply. This diversity is interpreted as a sign of quality.

Even if $\beta^{\text{up}}_{\text{CompAdv}}$ is negative and economically large in magnitude, it does not prove that the presence of Asian assemblers causally upgrades the quality of suppliers in Europe and North America. It could have been that the suppliers with high u^1 were of high quality before the creation of plants outside Asia by Asian assemblers. However, when combined with the evidence from the management literature about supplier development programs, it does seem as if some portion of supplier quality differences are due to the presence of the Asian assemblers.

A separate concern is that this approach treats u^1 as an economically exogenous characteristic, rather than recomputing the Asian market share for counterfactual sets of matches M in the right sides of sums of revenues inequalities. I have explored the specification where notationally u^1 is replaced by $z^{\text{up}}_{\text{AsianShare}}(M)$, which is recomputed for counterfactual sets of matches. The corresponding point estimate is $\beta^{\text{up}}_{\text{CompAdv}} = -0.01 \approx 0$ and the confidence intervals rule out economically large magnitudes. The reason is that a new effect is introduced to the model: the inequalities ask why more firms do not choose to supply parts to Asian assemblers if there is some quality upgrade from doing so? A reason outside of the model why this does not happen is the fixed cost of having an additional supplier participate in a supplier development program. Having explored an alternative, I return to the preferred specification, where a supplier's competitive advantage is an economically exogenous supplier characteristic.

Table 5 presents the point estimates from the preferred specification. The other covariates are the assembler and supplier specialization measures in Table 3, which have similar point estimates. The scale normalization is still $\beta_{\text{Cont.}} = \pm 1$, although with the interaction term also involving $z^{\text{up}}_{\text{Continent}}(M)$, the normalization can only be understood by substituting a typical value for u^1 into the new regressor, $z^{\text{up}}_{\text{CompAdv}}(M) = u^1 \cdot z^{\text{up}}_{\text{Continent}}(M)$.

The new addition to Table 5 is the estimate of $\beta^{\text{up}}_{\text{CompAdv}}$, which uses an estimate of the decrease in the importance of specialization at the continent-of-brand level for suppliers to Asian brands' assembly plants in Europe and North America as evidence that suppliers to Asian assemblers have higher quality. These suppliers can win business from non-Asian assemblers. The estimate of $\beta^{\text{up}}_{\text{CompAdv}}$ is -3.20 and the mean and standard deviation of u^1, not listed in the table, are 0.108 and 0.191, respectively. Therefore, a one-standard deviation change in u^1 creates a change of $-3.20 \cdot 0.19 = -0.601$ in the coefficient on the degree of specialization at the continent-of-brand level. A car parts supplier with a market share among Asian assemblers that is one standard deviation higher than the mean, a share of 0.30, will have a total coefficient on continent-of-brand specialization of $+1 - 3.20 \cdot 0.30 = 0.04$, or approximately 0. This is a large magnitude effect. The interpretation is that suppliers to Asian assemblers can go out and win business from non-Asian assemblers as well, but suppliers to European and North American assemblers cannot win as much business from suppliers from other continents. Thus, the evidence from sorting in the market for car parts suggests that domestic suppliers to assemblers with headquarters in Asia are in a unique competitive position, consistent with them having a quality advantage. While the cross-sectional empirical work alone cannot identify whether a quality increase causally occurred after the entry of Asian-based assemblers to Europe and North America, the estimates and the evidence from the management literature together suggest that having higher quality assemblers in Europe and North America raises the quality of suppliers. Thus, in the automotive industry there are indirect benefits to domestic suppliers and assemblers from the trade barriers that encourage Asian assemblers to locate in Europe and North America.

9 Conclusions

This paper introduces a new estimator for matching games and applies it to answer two policy questions surrounding the automotive industry. First, the paper estimates the relative loss in structural revenue to suppliers from decreased specialization from General Motors divesting Opel. A forced divestiture ends up hurting most suppliers only a little as the point estimates to the gains to specialization at the brand and market levels, which are not affected by the divestment, are higher than the gains to specialization at the parent group level. Second, the paper estimates the gain to, say, North American suppliers from the presence of Asian-based assemblers in North America. Suppliers to Asian assemblers have substantially more diverse portfolios of car parts, suggesting they are higher quality and can win business from firms of different origins. Both estimates are inferred from a new type of data, the equilibrium portfolios of car parts from each supplier.

In terms of methodology, this paper discusses the estimation of structural revenue functions in many-to-many matching games with transferable utility. These matching games allow endogenous transfers that are additively separable in payoffs. A pairwise stable equilibrium must satisfy sum of revenues inequalities: an exchange of one downstream firm each between two upstream firms cannot produce a higher sum of revenues.

I introduce a semiparametric maximum score estimator for matching games. A Monte Carlo for small and simple (one-to-one) matching markets shows that simulation estimators such as the method of moments and MCMC data augmentation maximum likelihood are quite slow. For matching games of any interesting size, maximum score is feasible while simulation estimators are not. The computational advantages of maximum score are that all nested integrals are avoided and all nested equilibrium calculations are omitted. Inequalities need to be included only with some positive probability, which is important given that the number of necessary conditions from pairwise stability increases rapidly with the number of agents in a matching market. Under additional assumptions, maximum score is consistent in the presence of multiple equilibria. The estimator uses data on only observed matches and agent characteristics. It does not require the often unavailable data on transfers, quotas, revenues or potential entrants.

A Proofs

A.1 Theorem 4.1: Consistency

Consistency follows from verifying the conditions of Theorem 2.1 in Newey and McFadden (1994).

A.1.1 The Limiting Objective Function is Globally Maximized at $\beta = \beta^0$

By a law of large numbers and the law of iterated expectations, the probability limit of the maximum score objective function is

$$Q_\infty(\beta) = \int_A \sum_{\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\}\in W(A)} C(\langle u_1,d_1\rangle,\langle u_2,d_2\rangle)\cdot 1\left[X_{u_1,u_2,d_1,d_2}{}'\beta \geq 0\right] d\rho(A),$$

where the dependence of ρ on the true model functions has been suppressed for conciseness.

For each pair of an assignment A_1 and a $\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\} \in W(A_1)$ in the integrand above, there is an assignment A_2 that is $A_2 = (A_1\backslash\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\})\cup\{\langle u_1,d_2\rangle,\langle u_2,d_1\rangle\}$ where $\{\langle u_1,d_2\rangle,\langle u_2,d_1\rangle\} \in W(A_2)$. By the construction of each X_{u_1,u_2,d_1,d_2}, the two inequalities are weakly exclusive: either $X_{u_1,u_2,d_1,d_2}{}'\beta > 0$ or $X_{u_1,u_2,d_2,d_1}{}'\beta > 0$, or they are equal. Because at least one element of X_{u_1,u_2,d_1,d_2} has continuous support, the probability that they are equal is 0 and does not contribute to the value of the integral. By an assumption to the theorem, $C(\langle u_1,d_1\rangle,\langle u_2,d_2\rangle) = C(\langle u_1,d_2\rangle,\langle u_2,d_1\rangle)$.

The ranking of the weights on the inequalities reduces to comparing $\rho(A_1)$ and $\rho(A_2)$. By the rank order property for many markets, $\rho(A_1) \geq \rho(A_2)$ whenever $X_{u_1,u_2,d_1,d_2}{}'\beta^0 \geq 0$. Therefore, $\beta = \beta^0$ will cause the higher of the two exclusive inequalities to enter the integrand for every pair of (A_1, A_2) and every $\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\} \in W(A_1)$. Therefore, the limiting objective function is globally maximized at $\beta = \beta^0$.

A.1.2 The Global Maximum $\beta = \beta^0$ Is Unique

Assume to the contrary, so that $\beta^1 \in \mathcal{B}$ gives $Q_\infty(\beta^1) = Q_\infty(\beta^0)$. The set of assignments and valid sum of revenues inequalities where β^0 and β^1 give different predictions for the rank ordering of $\rho(A_1)$

and $\rho(A_2)$ is

$$L\left(\beta^0,\beta^1\right)=\left\{A \mid \exists\left\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\right\}\in W(A) \text{ s.t. } X_{u_1,u_2,d_1,d_2}{}'\beta^0>0>X_{u_1,u_2,d_2,d_1}{}'\beta^1\right\}\cup \\ \left\{A \mid \exists\left\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\right\}\in W(A) \text{ s.t. } X_{u_1,u_2,d_1,d_2}{}'\beta^0<0<X_{u_1,u_2,d_2,d_1}{}'\beta^1\right\}.$$

If $\int_{A\in L(\beta^0,\beta^1)}\rho(A)\,dA>0$, then $H_\infty\left(\beta^1\right)<H_\infty\left(\beta^0\right)$ and β^0 will be the unique global maximizer. Recall that x_1 is one element of X_{u_1,u_2,d_1,d_2} with continuous support. Let $X^-_{u_1,u_2,d_1,d_2}$ be all other elements of X_{u_1,u_2,d_1,d_2} and let β_- be all other elements of β. Then write

$$L\left(\beta^0,\beta^1\right)=\left\{A \mid \exists\left\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\right\}\in W(A) \text{ s.t. } X^-_{u_1,u_2,d_1,d_2}{}'\beta^0_->-x_{1,u_1,u_2,d_2,d_1}>X^-_{u_1,u_2,d_1,d_2}{}'\beta^1_-\right\}\cup \\ \left\{A \mid \exists\left\{\langle u_1,d_1\rangle,\langle u_2,d_2\rangle\right\}\in W(A) \text{ s.t. } X^-_{u_1,u_2,d_1,d_2}{}'\beta^0_-<-x_{1,u_1,u_2,d_2,d_1}<X^-_{u_1,u_2,d_1,d_2}{}'\beta^1_-\right\},$$

where the coefficient on x_1 has been normalized to 1. The argument when the coefficient on x_1 is -1 is similar, as are the cases where $\beta^1_1=-1$ and $\beta^0_1=+1$ as well as $\beta^1_1=+1$ and $\beta^0_1=-1$. By an assumption to the theorem, x_{1,u_1,u_2,d_2,d_1} has support everywhere on the real line, so that $L\left(\beta^0,\beta^1\right)$ does have positive probability.

A.1.3 Continuity of the Limiting Objective Function and Uniform Convergence

Lemma 2.4 from Newey and McFadden (1994) can be used to prove continuity of $Q_\infty(\beta)$ as well as uniform-in-probability convergence of $Q_H(\beta)$ to $Q_\infty(\beta)$. Remember that the asymptotics are in the number of markets. The conditions of Lemma 2.4 are that the data (across markets) are i.i.d. (holds by assumption); that the parameter space $\mathcal{B}$ is compact (holds by assumption); that the terms for each market are continuous with probability 1 in β; and that the terms for each market are bounded by a function whose mean is not infinite. While the terms for each market are not continuous in β because of the indicator functions, they are continuous with probability 1 because each X_{u_1,u_2,d_1,d_2} has some elements with continuous support. The value of the objective function for a given market is bounded by the number of inequalities, which is finite.

A.2 Theorem 4.2

The limit of the objective function as $J\to\infty$ is

$$Q_\infty(\beta)=\int_{p_1,p_2}\sum_{d_1\in p_1}\sum_{d_2\in p_2}1\left[X_{u_1,u_2,d_1,d_2}{}'\beta\geq 0\right]g(p_1)\,g(p_2)\,dp_1dp_2,$$

where p_1 and p_2 are partner lists and g is the density over partner lists. Showing that $Q_\infty(\beta)$ is uniquely maximized at β^0 is very similar to the previous consistency proof and is therefore mostly omitted. Also, the technical requirements in Lemma 2.4 from Newey and McFadden (1994) are satisfied for the same reasons.

References

Abrevaya, Jason and Jian Huang, "On the bootstrap of the maximum score estimator," *Econometrica*, July 2005, *73* (4), 1175–1204.

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

Ahlin, Christian, "Matching for Credit: Risk and Diversification in Thai Microcredit Groups," December 2009. Michigan State University working paper.

Akkus, Oktay and Ali Hortacsu, "The Determinants of Bank Mergers: A Revealed Preference Analysis," March 2007. University of Chicago working paper.

Andrews, Donald W.K. and Marcia M.A. Schafgans, "Semiparametric Estimation of the Intercept of a Sample Selection Model," *Review of Economic Studies*, 1998, *65* (3), 497–517.

Baccara, Mariagiovanna, Ayse Imrohoroglu, Alistair Wilson, and Leeat Yariv, "A Field Study on Matching with Network Externalities," August 2009. New York University working paper.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, September 2007, *75* (5), 1331–1370.

_ , Han Hong, and Stephen Ryan, "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, September 2010, *78* (5), 1529–1568.

_ , Jeremy T. Fox, and Stephen Ryan, "Evaluating Wireless Carrier Consolidation Using Semiparametric Demand Estimation," *Quantitative Marketing and Economics*, 2008, *6* (4), 299–338.

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

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

Camiña, Ester, "A Generalized Assignment Game," *Mathematical Social Sciences*, 2006, *52* (2), 152–161.

Chernozhukov, Victor and Han Hong, "An MCMC Approach to Classical Estimation," *Journal of Econometrics*, 2003, *115* (2), 293–346.

_ , _ , and Elie Tamer, "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 2007, *75* (5), 1243–1284.

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

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, November 2009, *77* (6), 1791–1828.

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

Delgado, Miguel A., Juan M. Rodríguez-Poo, and Michael Wolf, "Subsampling inference in cube root asymptotics with an application to Manski's maximum score estimator," *Economics Letters*, 2001, *73*, 241–250.

Demange, Gabrielle, David Gale, and Marilda A. Oliveira Sotomayor, "Multi-Item Auctions," *The Journal of Political Economy*, August 1986, *94* (4), 863–872.

Dyer, Jeffrey H., "Specialized Supplier Networks as a Source of Competitive Advantage: Evidence from the Auto Industry," *Strategic Management Journal*, 1996, *17* (4), 271–291.

_ , "Effective interim collaboration: how firms minimize transaction costs and maximise transaction value," *Strategic Management Journal*, 1997, *18* (7), 535–556.

Echenique, Federico, "What Matchings Can Be Stable? The Testable Implications of Matching Theory," *Mathematics of Operations Research*, August 2008, *33* (3), 757–768.

Ekeland, Ivar, James J. Heckman, and Lars Nesheim, "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 2004, *112* (S1), 60–109.

Fox, Jeremy T., "Semiparametric estimation of multinomial discrete-choice models using a subset of choices," *RAND Journal of Economics*, Winter 2007, *38* (4), 1002–1019.

_ , "Identification in Matching Games," *Quantitative Economics*, 2010, *Forthcoming.*

_ **and Patrick Bajari**, "Measuring the Efficiency of an FCC Spectrum Auction," 2010. University of Michigan working paper.

Gale, David and Lloyd Shapley, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69.*

Garicano, Luis and Esteban Rossi-Hansberg, "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, November 2006, *121* (4), 1383–1435.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey, "Regular Quantal Response Equilibrium," *Experimental Economics*, December 2005, *8*, 347–367.

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

Haile, Philip A. and Elie Tamer, "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, February 2003, *111* (1), 1–51.

Han, Aaron K., "Nonparametric Analysis of a Generalized Regression Model: The Maximum Rank Correlation Estimator," *Journal of Econometrics*, 1987, *35*, 303–316.

Hatfield, John William and Paul R. Milgrom, "Matching with Contracts," *The American Economic Review*, September 2005, *95* (4), 913–935.

_ **and Scott Duke Kominers**, "Many-to-Many Matching with Contracts," February 2010. Stanford University working paper.

Hitsch, Günter, Ali Hortaçsu, and Dan Ariely, "Matching and Sorting in Online Dating Markets," *American Economic Review*, 2009.

Horowitz, Joel L., "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, May 1992, *60* (3), 505–551.

_ , "Bootstrap critical values for tests based on the smoothed maximum score estimator," *Journal of Econometrics*, 2002, *111*, 141–167.

Jacquier, E., M. Johannes, and N. Polson, "MCMC maximum likelihood for latent state models," *Journal of Econometrics*, 2007, *137* (2), 615–640.

Jaume, Daniel, Jordi Massó, and Alejandro Neme, "The Multiple-partners Assignment Game with Heterogeneous Sales and Multi-unit Demands: Competitive Equilibria," December 2009. Universidad Nacional de San Luis working paper.

Jun, Sung Jae, Joris Pinkse, and Yuanyuan Wan, "Cube-Root-N and Faster Convergence, Laplace Estimators, and Uniform Inference," 2009. Pennsylvania State University working paper.

Kamath, Rajan R. and Jeffrey K. Liker, "A Second Look at Japanese Product Development," *Harvard Business Review*, November-December 1994.

Kelso, Alexander S. and Vincent P. Crawford, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, November 1982, *50* (6), 1483–1504.

Kim, Jeankyung and David Pollard, "Cube Root Asymptotics," *The Annals of Statistics*, 1990, *18* (1), 191–219.

Klier, Thomas and James Rubenstein, *Who Really Made Your Car? Restructuring and Geographic Change in the Auto Industry*, Upjohn Institute, 2008.

Kojima, Fuhito, Parag A. Pathak, and Alvin E. Roth, "Matching with Couples: Stability and Incentives in Large Markets," April 2010. Stanford University working paper.

Kovalenkov, Alexander and Myrna Holtz Wooders, "Advances in the Theory of Large Cooperative Games and Application to Club Theory: the Side Payments Case," in Carlo Carraro, ed., *The Endogenous Formation of Economic Coalitions*, Edward Elgar: Chelenham, 2003.

Langfield-Smith, K. and M.R. Greenwood, "Developing Co-operative Buyer-Supplier Relationships: A Case Study of Toyota," *Journal of Management Studies*, 1998, *35* (3), 331–353.

Leonard, Herman B., "Elicitation of Honest Preferences for the Assignment of Individuals to Positions," *The Journal of Political Economy*, June 1983, *91* (3), 461–479.

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

Liker, Jeffrey K. and Yen-Chun Wu, "Japanese Automakers, U.S. Suppliers and Supply-Chain Superiority," *Sloan Management Review*, Fall 2000.

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

_ , "Semiparametric Analysis of Discrete Response," *Journal of Econometrics*, 1985, *27*, 313–333.

_ **and T. Scott Thompson**, "Operational Characteristics of Maximum Score Estimation," *Journal of Econometrics*, 1986, pp. 85–108.

Matzkin, Rosa L., "Nonparametric identification and estimation of polychotomous choice models," *Journal of Econometrics*, 1993, *58*, 137–168.

McFadden, Daniel L., "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics*, Academic Press, 1973.

Milgrom, Paul, "Putting Auction Theory to Work: The Simultaneous Ascending Auction," *Journal of Political Economy*, 2000, *108* (2), 245–272.

Mindruta, Denisa, "Value Creation in University-Firm Research Collaborations: A Matching Approach," February 2009. HEC Paris working paper.

Monteverde, Kirk and David J. Teece, "Supplier Switching Costs and Vertical Integration in the Automobile Industry," *The Bell Journal of Economics*, Spring 1982, *13* (1), 206–213.

Newey, Whitney K. and Daniel McFadden, "Large Sample Estimation and Hypothesis Testing," in "Handbook of Econometrics," Vol. IV, Elsevier, 1994, chapter 36, pp. 2111–2245.

Novak, Sharon and Birger Wernerfelt, "On the Grouping of Tasks into Firms: Make-or-Buy with Interdependent Parts," 2007. MIT working paper.

_ **and Scott Stern**, "How Does Outsourcing Affect Performance Dynamics? Evidence from the Automobile Industry," *Management Science*, December 2008, *54* (12), 1963–1979.

_ **and** _ , "Complementarity Among Vertical Integration Decisions: Evidence from Automobile Product Development," *Management Science*, 2009, *55* (2), 311–332.

_ **and Steven D. Eppinger**, "Sourcing by Design: Product Complexity and the Supply Chain," *Management Science*, 2001, *47* (1), 189–204.

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

Pakes, Ariel, "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 1986, *54* (4), 755–784.

—, **Jack Porter, Katherine Ho, and Joy Ishii**, "Moment Inequalities and Their Application," November 2006. Harvard University working paper.

Pinkse, C.A.P., "On the computation of semiparametric estimates in limited dependent variable models," *Journal of Econometrics*, 1993, pp. 185–205.

Politis, Dimitris N. and Joseph P. Romano, "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions," *The Annals of Statistics*, December 1994, *22* (4), 2031–2050.

—, —, **and Michael Wolf**, *Subsampling*, Springer, 1999.

Romano, J.P. and A.M. Shaikh, "Inference for the identified set in partially identified econometric models," *Econometrica*, 2010, *78* (1), 169–211.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 1974.

Roth, Alvin E. and Marilda A. Oliveira Sotomayor, *Two-Sided Matching* number 18. In 'Econometric Society Monographs.', Econometric Society, 1990.

Rust, John, "Optimal Replacement of the GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 1987.

Santiago, David and Jeremy T. Fox, "A Toolkit for Matching Maximum Score Estimation and Point and Set Identified Subsampling Inference," 2009. University of Michigan working paper.

Sherman, Robert P., "The Limiting Distribution of the Maximum Rank Correlation Estimation," *Econometrica*, January 1993, *61* (1), 123–137.

—, "Maximal Inequalities for Degenerate U-Processes with Applications to Optimization Estimators," *The Annals of Statistics*, March 1994, *22* (1), 439–459.

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

Sotomayor, Marilda, "The multiple partners game," in Mukul Majumdar, ed., *Equilibrium and dynamics: essays in honour of David Gale*, Macmillan, 1992, chapter 17, pp. 322–336.

Storn, Rainer and Kenneth Price, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, 1997, *115*, 341–359.

Subbotin, Viktor, "Asymptotic and bootstrap properties of rank regressions," November 2007. Northwestern University working paper.

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

Figure 1: GMM Run Time Is Convex In the Number of Firms on One Side of a Matching Market

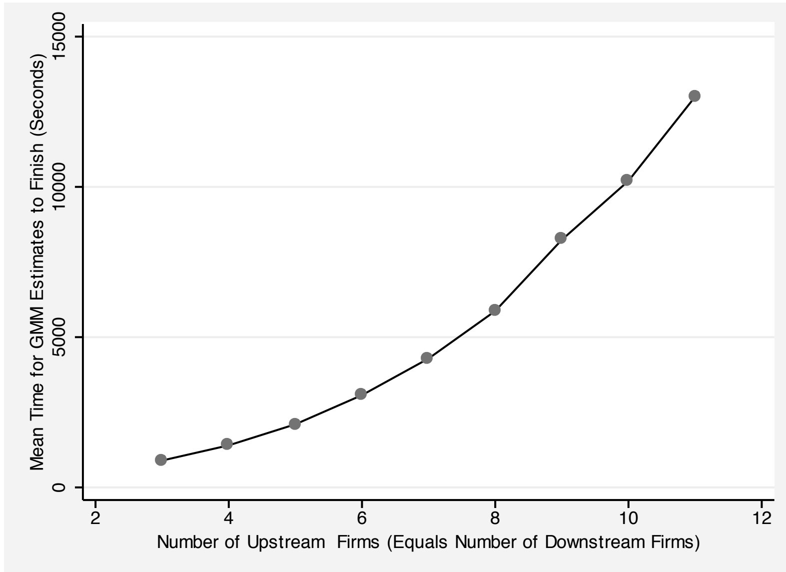

Table 1: Monte Carlo Results Showing Maximum Score Bias and RMSE Decrease with Sample Sizes, True Value $\beta_2 = 1.5$

# Upstr. & # Downstr.	# Markets	Errors Std. Dev.	Bias	RMSE	# Upstr. & # Downstr.	# Markets	Errors Std. Dev.	Bias	RMSE
3	100	1	-0.12	0.66	3	100	10	-0.44	1.03
10	10	1	-0.07	0.59	10	10	10	-0.41	1.04
30	10	1	0.03	0.37	30	10	10	-0.08	0.76
60	10	1	0.04	0.27	60	10	10	0.01	0.64
60	40	1	0.01	0.16	60	40	10	0.03	0.47
100	1	1	0.05	0.43	100	1	10	-0.19	0.88
200	1	1	0.02	0.31	200	1	10	-0.08	0.79

The true parameter is $\beta_2 = 1.5$. The population bias is $E\left[\hat{\beta}_2 - 1.5\right]$ and the population RMSE is $\sqrt{E\left[\left(\hat{\beta}_2 - 1.5\right)^2\right]}$, where 1.5 is the value of β_2 used to generate the fake data. The model is estimated 1000 times for each experiment. A fake dataset consists of the listed number of independent markets. New observable variables X and match-specific errors of the form $\epsilon_{\langle u,d\rangle}$ are drawn for each market and each replication. Each market is a one-to-one, two-sided matching game. The number of upstream firms always equals the number of downstream firms. The equilibrium assignment is calculated using a linear programming problem. In the left panel, the match-specific errors have the mixed normal distribution $0.4 \cdot N\left(0, 2^2\right) + 0.6 \cdot N\left(5, 1^2\right)$, which has a standard deviation of 1. This is a bimodal density. In the right panel, the error distribution is $0.4 \cdot N\left(0, 20^2\right) + 0.6 \cdot N\left(5, 10^2\right)$, which has a standard deviation of 10.

Each agent has a vector of two types, each drawn from a bivariate normal with means of 1, standard deviations of 1, and covariances of 1/2. The coefficient on the product of the first types is normalized to one. The estimate of the sign of the coefficient is superconsistent and so I do not explore its finite sample properties. The value of being unmatched is zero and unmatched firms are included in the analysis, in order to facilitate comparisons with parametric estimators in Table 2. The parametric estimators are inconsistent if unmatched firms are included in the true model but are not in the data and the equilibrium is computed as if those firms are not present.

Table 2: Comparing Maximum Score to Parametric Estimators, True Value $\beta_2 = 1.5$

			Maximum Score			Likelihood MCMC			Method of Moments		
True Distribution	Errors Std. Dev.	Number of Upstream Firms	Bias	RMSE	Time (s)	Bias	RMSE	Time (s)	Bias	RMSE	Time (s)
Normal	1	3	-0.16	0.54	2	-0.13	0.19	2700	0.03	0.337	1400
Normal	25	3	-0.54	1.14	2	0.28	0.74	2700	-0.26	2.00	1400
Normal	50	3	-0.37	1.14	2	0.49	1.86	2900	-0.14	3.54	1500
Normal	1	4	-0.02	0.50	2	-0.16	0.20	7100	0.01	0.31	2100
Normal	25	4	-0.42	1.13	2	-0.07	0.79	7000	-0.03	1.59	2000
Normal	50	4	-0.77	1.30	2	-0.17	1.55	7000	-0.37	2.61	2100
Mixed Normal	1	3	-0.11	0.69	2	-1.19	1.19	2700	-1.09	1.10	1400
Mixed Normal	25	3	-0.46	1.17	1	-0.31	1.12	2400	-0.58	1.96	1000
Mixed Normal	50	3	-0.60	1.46	1	-0.33	2.27	2400	0.16	3.77	1300
Mixed Normal	1	4	0.05	0.65	2	-1.13	1.13	7100	-1.07	1.07	2100
Mixed Normal	25	4	-0.68	1.15	2	-0.47	0.84	7000	-0.56	1.32	2000
Mixed Normal	50	4	-0.93	1.58	2	-0.16	1.26	7200	-0.57	2.61	2100

The true parameter is $\beta_2 = 1.5$. The population bias is $E\left[\hat{\beta}_2 - 1.5\right]$, and the population RMSE is $\sqrt{E\left[\left(\hat{\beta}_2 - 1.5\right)^2\right]}$, where 1.5 is the value of β_2 used to generate the fake data. The model is estimated 100 times for each experiment. The same fake dataset is used for all three estimators. A fake dataset consists of 100 independent matching markets. New observable variables X and match-specific errors of the form $\epsilon_{\langle u,d \rangle}$ are drawn for each market and each replication. Each market is a one-to-one, two-sided matching game. The number of upstream firms always equals the number of downstream firms. The equilibrium assignment is calculated using a linear programming problem. Each firm has a vector of two types, each drawn from a bivariate normal with means of 1, standard deviations of 1, and covariances of 1/2. In the top panel, the match-specific errors $\epsilon_{\langle u,d \rangle}$ have $N\left(0, \sigma^2\right)$ distributions, where σ is the standard deviation listed in the table. In the bottom panel, the errors have the mixed normal distribution $0.4 \cdot N\left(-6, \sigma_1^2\right) + 0.6 \cdot N\left(4, \sigma_2^2\right)$, where the total dispersion of $\epsilon_{\langle u,d \rangle}$ is given in the table. This is a bimodal density.

Each agent has a vector of two types. For maximum score, the coefficient on the product of the first types is normalized to one. The estimate of the sign of the coefficient is superconsistent and so I do not explore its finite sample properties. Coefficients on both parameters are estimated for the parametric estimators, as the scale normalization is on the error term. The coefficient estimates of the parametric estimators are multiplied by the true standard deviation of the errors, to make the scale normalization comparable to maximum score. All experiments were done on the same computer using MATLAB. Times listed are CPU times.

Table 3: Specialization By Suppliers and Assemblers

	Revenue Function Estimates		Sample Statistics for HHI Measures	
HHI Measure	Point Estimate	95% CI	Mean	Standard Deviation
Suppliers				
Continent	+1	Superconsistent	0.725	0.210
Parent Group	4.09	(1.58,5.76)	0.395	0.291
Brand	11.6	(3.37,18.9)	0.283	0.292
Model	86.0	(75.9,132)	0.201	0.286
Assemblers				
Parent Group	-0.0698	(-0.106,-0.013)	0.312	0.196
Brand	-0.0294	(-4.37,0.008)	0.449	0.204
Model	-0.0315	(-7.38,0.254)	0.786	0.162
# Inequalities	332,218			
% Satisfied	80.6%			

I estimate $\beta^{\text{up}}_{\text{Cont.}}$ by optimizing the maximum score objective function over the other parameters, first fixing $\beta_{\text{Cont.}} = +1$ and then fixing $\beta_{\text{Cont.}} = -1$. I then take the set of estimates corresponding to the maximum of the two objective function values as the final set of estimates. The estimate of a parameter that can take only two values is superconsistent, so I do not report a confidence interval.

I use the numerical optimization routine differential evolution, in Mathematica. For differential evolution, I use a population of 200 points and a scaling factor of 0.5. The numerical optimization is run 30 times with different initial populations of 200 points. I take the point estimates corresponding to the maximum reported objective function value over the 30 runs. For inference, I use subsample sizes equal to 1/4 of the matching markets. Unfortunately, the literature on subsampling has not produced data dependent guidelines for choosing the subsample size. I use 150 replications (artificial datasets) in subsampling. Following the asymptotic theory, I sample from the 187 distinct matching markets (component categories). Constructing the inequalities, producing the 30 estimates, and constructing confidence intervals took 9.5 hours on a single core of a 2010 vintage desktop computer.

In the maximum score objective function, an inequality is satisfied if the left side exceeds the right side by 0.0001. This small perturbation to the sum of revenues on the right side ensures that inequalities such as $0 > 0$ will not be counted as being satisfied because of some numerical-approximation error resulting in, say, $2.0 \times 10^{-15} > 1.0 \times 10^{-15}$.

Table 4: Percentage Revenue Change By Suppliers From GM Divesting Opel

Quantile	
0	-0.090
0.10	-0.028
0.25	-0.016
0.50 (median)	-0.007
0.75	-0.003
0.90	-0.001
1	~0

This table uses the point estimates from Table 3 to calculate the structural revenues of suppliers before and after GM divests Opel. In the model, Opel becomes a separate parent group. For each firm selling one or more parts to Opel and one or more cars to another GM brand, I calculate $\frac{\beta^{\text{up}}_{\text{PG}} \Delta z^{\text{up}}_{\text{ParentGroup}}(M_u)}{r^{\text{up}}_{\beta^{\text{up}}}(M_u)}$.

Table 5: Supplier Competitive Advantages From Asian Assemblers

	Revenue Function Estimates		Sample Statistics for HHI Measures	
HHI Measure	Point Estimate	95% CI	Mean	Standard Deviation
Suppliers				
Continent	+1	Superconsistent	0.725	0.210
Parent Group	4.04	(2.29,5.49)	0.395	0.291
Brand	11.6	(6.68,18.8)	0.283	0.292
Model	84.3	(74.3,127)	0.201	0.286
Competitive Advantage	-3.20	(-4.71,-2.49)	0.069	0.160
Assemblers				
Parent Group	-0.0940	(-0.147,-0.048)	0.312	0.196
Brand	-0.0219	(-2.95,0.006)	0.449	0.204
Model	-0.0381	(-5.08,0.216)	0.786	0.162
# Inequalities	332,218			
% Satisfied	80.8%			

See Table 3 for implementation details.