cemmap

centre for microdata methods and practice

Estimating Matching Games with Transfers

Jeremy T. Fox

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP14/16

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Estimating Matching Games with Transfers

Jeremy T. Fox*
Rice University and NBER

March 2016

Abstract

I explore the estimation of transferable utility matching games, encompassing many-to-many matching, marriage and matching with trading networks (trades). I introduce a matching maximum score estimator that does not suffer from a computational curse of dimensionality in the number of agents in a matching market. I apply the estimator to data on the car parts supplied by automotive suppliers to estimate the returns from different portfolios of parts to suppliers and automotive assemblers.

*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 and workshop participants at many universities and conferences. Chenchuan Li, David Santiago, Louis Serranito and Chenyu Yang provided excellent research assistance. Email: jeremyfox@gmail.com.

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 valuations of agents over the characteristics of potential partners and other measured aspects of the relationships. This is a challenging task compared to estimating valuations using more traditional data because we observe only the equilibrium relationships and not each agent's equilibrium choice set: the identity of the other agents who would be willing to match with a particular agent. We must infer utility parameters from the sorting seen in the data.

This paper presents an estimator for transferable utility matching games. Transferable utility matching games feature prices (or transfers) for relationships. This paper primarily models the formation of relationships as a competitive equilibrium to the matching model of Azevedo and Hatfield (2015, Section 6), which uses a continuum of agents. This model is quite general and includes many special cases of empirical interest.[1] Equilibrium existence and uniqueness are generically satisfied. The paper briefly considers the related model of Hatfield, Kominers, Nichifor, Ostrovsky and Westkamp (2013), which uses a finite number of agents. In these models, a generalization of a match is called a trade. A trade can include other aspects in addition to the identity of the match partners. For example, in a labor market a trade could specify the number of hours an employee is to work per week and the number of vacation days per year.

Using this structure, the paper explores the estimation of valuation functions, which represent the structural preference of agents for matches or trades. Computational challenges are key in matching and a computationally simple matching maximum score estimator is introduced to address the computational challenges. In an application, the paper uses the matching maximum score estimator to empirically answer questions related to the car parts industry. I first describe the methodological contribution and then the empirical application.

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 agents in them, compared to, say, the two to five agents 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 so-called car component category. In Fox and Bajari (2013), we apply the estimator introduced 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 car parts and auction datasets are rich. There is a lot of information on agent characteristics and unknown parameters that can be learned from the observed sorting of car part suppliers to car assemblers or bidders to items for sale. To take advantage of rich data sets, a researcher must use an estimator that allows for the dimensionality of typical problems.

The solution in maximum score is to introduce inequalities that are computationally simple to work with. The objective function is proportional to the number of inequalities that are true at a

[1] Some initial papers on one-to-one, two-sided matching with transferable utility are Koopmans and Beckmann (1957); Gale (1960); Shapley and Shubik (1972); Becker (1973). 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.

value for the unknown parameters in the valuation function. The inequalities involve only observable characteristics of matches, trades and unknown parameters. While there are unobservables in the true matching model, the maximum score inequalities do not require numerically integrating out unobservables, as in simulation estimators (McFadden, 1989; Pakes and Pollard, 1989; Hajivassiliou and McFadden, 1998). Therefore, the estimator does not suffer from computational issues due to numerical integration. Further, the integrand of the integral in a simulation estimator for a game theoretic model often involves a nested fixed point procedure to compute an equilibrium to the game (for each simulation draw). No nested fixed point procedure is used in maximum score. Finally, only a subset of the valid inequalities in maximum score can be included without losing the estimator's (point or set) consistency.

The curse of dimensionality for other estimators can be seen for a simple, one-to-one matching market with 100 men and 100 women. If no agents can be unmatched, for simplicity, there are $100^2 = 10{,}000$ possible matches and $100! \approx 10^{158}$ possible assignments. If there are match-specific unobservables, the likelihood function involves a numerical integral of 10,000 dimensions and an integrand that involves computing whether 1 out of 10^{158} assignments occur. Of course, simulation estimators (sometimes based on moment conditions chosen for tractability) can be used for matching when required at least when markets are small and computational resources are large, as in Boyd, Lankford, Loeb and Wyckoff (2013), Sørensen (2007), Agarwal and Diamond (2013) and Fox, Hsu and Yang (2015).

The computational advantages of matching maximum score are also present in single agent, multinomial choice maximum score (Manski, 1975; Matzkin, 1993; Briesch, Chintagunta and Matzkin, 2002; Fox, 2007). Indeed, this paper shows that, under some assumptions, a single agent maximum score estimator can be applied to matching data if the prices of matches or trades are observed in the data. Akkus, Cookson and Hortacsu (forthcoming) and Fox and Bajari (2013, Section VI.B, Appendix C) apply single agent maximum score estimators to matching data with prices. However, in many applications prices are not observed. For example, in marriage prices paid between men and women are rarely observed and prices are private contractual details in the car parts industry studied here. Without price data, a contribution of this paper is to propose matching maximum score inequalities that do not use price data. The paper proves that the maximum score setup results at least in set identification under the Azevedo and Hatfield (2015, Section 6) model with a continuum of agents and other assumptions.

The use of a model with a continuum of agents as the true or limiting matching model dates to the pioneering work on estimating matching games by Choo and Siow (2006). Choo and Siow studied the case of one-to-one, two-sided matching or marriage. They assume that the unobservables have the type I extreme value distribution, resulting in a logit choice model at the agent level and closed form formulas for matching patterns.[2] Chiappori, Salanié and Weiss (2015) and Galichon and Salanie (2012) use the term separability to highlight a key assumption in Choo and Siow. Separability restricts how agents' unobservable components of valuations vary across matches or trades. The main results of this

[2]Dagsvik (2000) provides logit-based methods for studying matching games where other aspects of a relationship than money are also part of the pairwise stable 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 estimators typically use data on the prices of trades (Rosen, 1974; Ekeland, Heckman and Nesheim, 2004; Heckman, Matzkin and Nesheim, 2010).

paper also rely on separability and this assumption is discussed in some detail below. Separability fits well into the matching model of Azevedo and Hatfield (2015, Section 6). Previous versions of this paper introduced the matching maximum score estimator for many-to-many, two-sided matching and used the closed form logit formulas from Choo and Siow to show the estimator's consistency for the simpler marriage model. However, it was Graham (2011, Theorem 4.1) who, in a survey article discussing previous drafts of this paper, first proved set identification for the continuum marriage model under semiparametric conditions nearly identical to those used for proving set identification for single agent maximum score models (Manski, 1975; Matzkin, 1993; Briesch, Chintagunta and Matzkin, 2002; Fox, 2007). The current version of this paper introduces the matching maximum score estimator, uses the setup of the quite general Azevedo and Hatfield (2015, Section 6) model and extends the argument of Graham for marriage to the more general setting.

This paper was originally part of a larger project including Fox (2007) on maximum score methods for single agent multinomial choice, Fox (2010) on nonparametric identification in matching games, and Fox and Bajari (2013) on an empirical application of the matching maximum score estimator to an FCC spectrum auction. Earlier working paper versions of this paper have circulated since 2005, at the same time as drafts of these other papers. The methodological contribution of the current paper is the introduction of the matching maximum score estimator and the discussion of its theoretical (set) consistency using the high generality of the matching model of Azevedo and Hatfield (2015, Section 6). Earlier drafts of this paper focused on models of many-to-many, two-sided matching and special cases thereof (such as one-to-one matching). The empirical work on car parts is also novel.

The matching maximum score estimator has been used in empirical work in many published papers and circulating working papers by now. Citing many of them would risk offending those omitted, but the breadth of applications to different fields of applied microeconomics and business shows the usefulness of the maximum score approach to matching estimation. Graham (2011) and Chiappori and Salanié (forthcoming), both mostly for marriage, as well as Mindruta, Moeen and Agarwal (2015), for the academic field of strategy, are three published surveys that extensively discuss the matching maximum score estimator. The Mindruta et al. survey is particularly useful for those wishing to compare a matching game theoretic approach to working with relationship data to other empirical approaches, such as so-called dyad regressions. I have made code for the matching maximum score estimator available (Santiago and Fox, 2009).

I cite some other methodological papers on matching in the rest of the text. One paper that is not cited below is Menzel (2015), who studies marriage and shows that a class of semiparametric non-transferable utility matching models (not considered here) converge, as the market grows large, to a parametric matching model with matching formulas quite similar to the matching formulas in the logit transferable utility marriage matching model with a continuum of agents in Choo and Siow (2006). The current paper consider the continuum matching model of Azevedo and Hatfield (2015, Section 6) and the semiparametric assumptions are made in the continuum; in no sense is a parametric model used as an approximation to the continuum.

1.1 Empirical Application to Car Parts

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 car parts industry. The first question relates to the productivity loss to suppliers from breaking up large assemblers of cars. During a recent large recession, North American-based automobile assemblers went through financial distress. As a consequence, North American-based assemblers divested or closed both North American brands (General Motor's Saturn) and European brands (Ford's Volvo) and 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 Opel, which was a serious policy debate in Germany in 2009, then any benefit from coordinating 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 estimate the valuations 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 place this investigation in an appendix as the key parameter has a wide confidence region.

I answer both of the above questions using the identities of the companies that supply each car part. The data list each car model and each car part on that model, and importantly the supplier of each car part. The intuition is that the portfolio of car parts that each supplier manufactures informs us 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 quality advantage that those suppliers have.

This paper takes the stand that the sorting pattern of sellers (suppliers like Bosch and Delphi) to buyers (assemblers like General Motors and Toyota) informs us about so-called valuation functions, key components of total profits, generating the payoffs of particular portfolios of car part trades to assemblers and to suppliers. In turn, the valuation functions for assemblers and suppliers help us answer the policy questions about government-induced divestitures and foreign assembler plants in North America. The loss to a supplier from GM divesting Opel occurs when supplying two car parts to a large parent company generates more valuation than supplying one car part each to two different assemblers. Thus, the valuation of a portfolio of trades is not necessarily the sum of the valuations from individual trades. The portfolio of trades of each firm is critical for valuations. Therefore, valuation functions are not additively separable across multiple trades, as they are in most prior work on matching not employing the maximum score estimator introduced in this paper.

I model the market for car parts as a two-sided, many-to-many matching game, with the two sides being assemblers and suppliers. In a competitive equilibrium, each firm will form the trades, car part transactions, that maximize its profits at the market-clearing prices. However, those prices are confidential contractual details not released to researchers. This motivates the use of the matching maximum score estimator without prices.

2 Example 1: Matching Maximum Score Estimator for Marriage

I first introduce the matching maximum score estimator for the simple case of one-to-one, two-sided matching, which I label Example 1. While many examples fit a one-to-one, two-sided matching setup, the case getting the most attention in the structural empirical literature is heterosexual, monogamous marriage. In order to communicate the implementation of the proposed method to empirical researchers, this section discusses only the estimator that a researcher codes on the computer and does not cover many underlying modeling details, such as unobservable terms. These modeling details are introduced in Section 3.

The researcher has data on $\tilde{N}$ marriages. View these as a random sample of the marriages from a larger market. For each marriage, the researcher observes some characteristics of the male j_m and some characteristics of the female j_f. The researcher uses these characteristics to specify a **joint valuation function based on observable types** $\tilde{X}(j_m, j_f)' \theta$, which is specified to be equal to a finite vector of observables $\tilde{X}(j_m, j_f)$ times a vector of unknown parameters θ. The vector $\tilde{X}(j_m, j_f)$ is chosen by the researcher based on the available data on agent characteristics. The choice of the elements of $\tilde{X}(j_m, j_f)$ ideally should be informed by results on the nonparametric identification of joint valuation functions in matching games, such as the discussion of Becker (1973) and Fox (2010) below. Note that the researcher can construct the vector $\tilde{X}(j_m, j_f)$ for a marriage that did not form in the data by varying the observable characteristics of the male and female.

The underlying transferable utility matching model follows Becker (1973) and the large subsequent literature and uses prices for marriage. Men pay women possibly negative transfers. However, consistent with actual data and previous structural empirical work on transferable utility matching models of marriage, prices are assumed not to be in the available data (Choo and Siow, 2006). Maximum score estimation with price data is discussed below.

Let i index the male in a marriage and let $f(i)$ index the female married to him in equilibrium. So j_m^i is the characteristics of male i and $j_f^{f(i)}$ is the characteristics of female $f(i)$. A competitive equilibrium (equivalent in this model to a stable matching) to this transferable utility matching game is **efficient**: it maximizes the sum of joint valuations across all marriages in the entire economy. If there were no unobservables (and recall that there are in the true model), an implication of social efficiency is that swapping spouses lowers the sum of joint valuations. Therefore, intuition suggests the following **matching maximum score inequality** should be useful in estimation:

$$\tilde{X}\left(j_m^{i_1}, j_f^{f(i_1)}\right)'\theta + \tilde{X}\left(j_m^{i_2}, j_f^{f(i_2)}\right)'\theta \geq \tilde{X}\left(j_m^{i_1}, j_f^{f(i_2)}\right)'\theta + \tilde{X}\left(j_m^{i_2}, j_f^{f(i_1)}\right)'\theta. \tag{1}$$

The inequality says that the sum of joint valuations of two observed matches should weakly exceed the sum of the joint valuations from the two matches where the married couples exchange spouses: male i_1 is married to female $f(i_2)$ instead of his actual wife $f(i_1)$, for example. If single people are in the data, one can normalize the valuation from being single to 0, add a constant term to $X(j_m, j_f)$, and include being single as a "spouse" for males and females in matching maximum score inequalities. Using data on singles would slightly complicate the notation for the summations below. Note that the competitive equilibrium being efficient is used for intuition for motivating matching maximum score inequalities in the main text. However, the formal proofs in the appendix do not rely on a competitive equilibrium being efficient.

The **matching maximum score objective function** is

$$\sum_{i_1=1}^{\tilde{N}-1}\sum_{i_2=i_1+1}^{\tilde{N}} 1\left[\tilde{X}\left(j_m^{i_1}, j_f^{f(i_1)}\right)'\theta + \tilde{X}\left(j_m^{i_2}, j_f^{f(i_2)}\right)'\theta \geq \tilde{X}\left(j_m^{i_1}, j_f^{f(i_2)}\right)'\theta + \tilde{X}\left(j_m^{i_2}, j_f^{f(i_1)}\right)'\theta\right]. \tag{2}$$

The dependent variable here is the matches in the data, $f(i)$ for man i. The matching maximum score, or maximum rank correlation as explained below, objective function checks whether each matching maximum score inequality is true. If an inequality is true for a guess of the parameter vector θ, the objective function increases by 1. Not all inequalities will be true even at the true value of the parameter vector θ because of unobservable variables present in the model, as introduced below. The objective function is a step function. The code in Santiago and Fox (2009) uses the differential evolution global optimization routine; other choices are possible. If the parameter θ in the maximum score model is point identified (as discussed below), any maximizer of the objective function is a consistent estimator. If the underlying parameter θ in the model is instead set identified, then a valid 95% confidence set for the identified parameters should be reported. Note that point and set identification as used here are properties of the probability limit (expectation) of the objective function, not the objective function for a finite sample of data. More discussion of inference and point vs set identification is below.

The matching maximum score objective function is easy to compute: it involves only addition, multiplication and checking an inequality. There is no numerical integration over unobservables and no nested fixed point computation of equilibria. There is no attempt to estimate a distribution of unobservables using sieve methods or to estimate matching probabilities nonparametrically in a first stage. Further, the objective function is written as if all possible inequalities will be used. But the available inequalities can be randomly sampled and the estimator will still be consistent. Altogether, I can say that the estimator does not suffer from a computational or data curse of dimensionality in the size of the data on the matching market in question.

If the researcher has data on D independent matching markets (say different towns in the marriage example), then the researcher merely adds an extra summation (and market indices d) to the objective function

$$\sum_{d=1}^{D}\sum_{i_1=1}^{\tilde{N}_d-1}\sum_{i_2=i_1+1}^{\tilde{N}_d} 1\left[X\left(j_m^{i_1,d}, j_f^{f(i_1,d)}\right)'\theta + X\left(j_m^{i_2,d}, j_f^{f(i_2,d)}\right)'\theta \geq X\left(j_m^{i_1,d}, j_f^{f(i_2,d)}\right)'\theta + X\left(j_m^{i_2,d}, j_f^{f(i_1,d)}\right)'\theta\right]. \tag{3}$$

Asymptotics could be in the number of markets D or, for fixed D, some notion $\tilde{N}$ of the number of recorded marriages such that each $\tilde{N}_d = \nu_d \cdot \tilde{N}$ for the fixed-with-$\tilde{N}$ market-specific proportionality factors ν_d, as discussed below.

3 Matching Game

Matching games model relationship formation. This paper discusses estimation of the transferable utility matching game with a continuum of agents in Azevedo and Hatfield (2015, Section 6), henceforth abbreviated as the AH model. An otherwise similar model with a finite number of agents in is Hatfield, Kominers, Nichifor, Ostrovsky and Westkamp (2013). Valuations in the AH model encompass valuations for a great many applications of empirical interest, including the previous example of marriage and many-to-many, two-sided matching with valuations defined over sets of matches. This generality is useful for scholars applying the matching maximum score estimator to matching games other than marriage and the one estimated for the car parts industry later in this paper. I discuss several special cases of the AH model below. The AH model has desirable properties: a unique competitive equilibrium exists under fairly innocuous technical conditions.

The arguments in this section lead to an estimation approach that uses data on one large matching market, as considered previously for one-to-one matching by Choo and Siow (2006) and the related work cited in the introduction. The asymptotic argument increases the sample size as the number of agents observed in the data grow large. In an early application, Fox and Bajari (2013) use the matching maximum score estimator and the large market asymptotic argument to study a large spectrum auction.

In the large market asymptotic argument, the limiting matching game, here the AH model, has a continuum of agents. AH show that there exists a unique equilibrium to this matching game. The equilibrium is deterministic in the aggregate. As a researcher collects more data, the asymptotic fiction is that the researcher is observing more agents from this limiting game. Therefore, the asymptotic fiction of collecting more data does not alter the outcome of the matching game in question; the researcher is merely learning more about an existing market with a continuum of agents.

3.1 The AH Model

For compatibility with Azevedo and Hatfield (2015, Section 6), I borrow much of the terminology and the notation of the AH model. I first lay out the general model and then discuss examples below. Let there be a set of **full agent types** I and a *finite* set of **trades** Ω. An agent of type $i \in I$ has a **valuation function** $v^i(\Phi, \Psi)$, where $\Phi \subseteq \Omega$ is the set of trades for which agent i is a **buyer** and $\Psi \subseteq \Omega$ is the set of trades for which agent i is a **seller**. The valuation function $v^i(\Phi, \Psi)$ takes on values in $[-\infty, \infty)$. The empty set $\emptyset$ refers to making no trades as, say, a buyer; $v^i(\emptyset, \emptyset)$ is normalized to

0.

Consider a **price** p_ω for each trade $\omega \in \Omega$. Let $p^\Omega = (p_\omega)_{\omega\in\Omega}$ be the **price vector** for all trades $\omega \in \Omega$. Under transferable utility, the **profit** of an agent i who buys trades Φ and sells trades Ψ at the prices p^Ω is

$$v^i(\Phi,\Psi) - \sum_{\omega\in\Phi} p_\omega + \sum_{\omega\in\Psi} p_\omega. \tag{4}$$

As in AH, there is a measure $\eta(i)$ over the set of agent types $i \in I$.[3] An **allocation** A is a map from the set of agent types I to the space of distributions over the product space formed by two power sets of Ω,

$$\mathcal{P}(\Omega) \times \mathcal{P}(\Omega).$$

For each type $i \in I$, the allocation A specifies a distribution $A^i(\Phi,\Psi)$ over sets of trades as a buyer Φ and as a seller Ψ; each $A^i(\Phi,\Psi)$ is the fraction of agents of full type i that conduct the trades Φ and Ψ.

An **arrangement** $\left(A, p^\Omega\right)$ is comprised of an allocation A and a price vector p^Ω. The allocation A is **incentive compatible** given the price vector p^Ω if each agent maximizes its profits (4) in the sense that $A^i(\Phi,\Psi) > 0$ only if

$$(\Phi,\Psi) \in \arg\max_{\tilde{\Phi}\subseteq\Omega,\tilde{\Psi}\subseteq\Omega} \left(v^i\left(\tilde{\Phi},\tilde{\Psi}\right) - \sum_{\omega\in\tilde{\Phi}} p_\omega + \sum_{\omega\in\tilde{\Psi}} p_\omega \right).$$

The allocation A is **feasible** if the **excess demand** for each trade $\omega \in \Omega$,

$$\int_I \left(\sum_{\Phi\supseteq\{\omega\}} \sum_{\Psi} A^i(\Phi,\Psi) - \sum_{\Psi\supseteq\{\omega\}} \sum_{\Phi} A^i(\Phi,\Psi) \right) d\eta(i), \tag{5}$$

equals 0. In the definition of excess demand, the sums are over subsets of the finite set of trades Ω and $\Phi \supseteq \{\omega\}$ means sum over sets where the trade ω is an element. The arrangement (A,p) is a **competitive equilibrium** if the allocation A is incentive compatible given the price vector p^Ω and is feasible.[4]

AH prove that a competitive equilibrium exists and is efficient in the sense of the allocation component A being the supremum of the social welfare function

$$\int_I \left(\sum_{\Phi} \sum_{\Psi} v^i(\Phi,\Psi)\, A^i(\Phi,\Psi) \right) d\eta(i), \tag{6}$$

where the supremum is taken over feasible allocations A. If the distribution of full types $\eta(i)$ ensures uniqueness of the maximizer of the social welfare function, then there is a unique allocation A corresponding with a competitive equilibrium. Further, AH prove that if $\eta(i)$ has full support in a precise

[3] I place the technical conditions from AH in footnotes. The measure η is defined with respect to some $\sigma-$algebra and satisfies $\eta(I) < \infty$. Further, the valuation function v^i must be a measurable function of i.

[4] AH require two further technical conditions. Using their words and skipping their notation: 1) The integral of absolute values of utility is finite as long as agents are not given bundles for which they have utility of $-\infty$ and 2) agents can supply any sufficiently small net demand for trades.

sense then the price vector p^{Ω} in the competitive equilibrium is unique. Given these relatively weak conditions for equilibrium existence and uniqueness, I maintain existence and uniqueness in what follows. What actually matters for the empirical approach is that the same competitive equilibrium is being played by all agents in the continuum market.

3.2 Observable and Unobservable Types

The AH model has complete information in the sense that no attributes of a trade are privately observed. Still, econometricians wish to distinguish between attributes of agents measured in the data and attributes not measured in the data. Recall i indexes a full agent type in the AH model. Let j index an **observable agent type** in a *finite* set of observable agent types J and let k index an **unobservable agent type** in some, likely infinite, set K, so that each full agent type i uniquely corresponds to a pair of agent types (j, k). This focus on observables and unobservable (in data) agent types specializes the AH model but does not alter its properties mentioned before.

The set of trades Ω is finite in the AH model and here the set of observable agent types J is also finite. Assume that the definition of a trade ω encodes the observable agent type $b(\omega) \in J$ for the buyer on trade ω and the observable agent type $s(\omega) \in J$ for the seller on trade ω. In the AH model, agents have preferences defined over trades. So an agent $i \in I$ is allowed to have preferences over the observable agent types $j \in J$ of the counterparties to trades ω. In some examples of the AH model below, a trade ω encodes *only* the observable agent types of the buyer $b(\omega)$ and seller $s(\omega)$. Trades may encode other aspects, as also discussed below.

Trades ω do not encode unobservable types k. So agents are not allowed to have preferences over the unobservable types $k \in K$ of counterparties on trades ω. The set of unobservable (in data) agent types K is not required to be finite. Indeed, a finite set of observable agent types J and an uncountably infinite set of unobservable agent types K in a model with a continuum of agents was first used in a parametric empirical model for one-to-one, two-sided matching (specifically, marriage) by Choo and Siow (2006). Later for one-to-one, two-sided matching, Galichon and Salanie (2012) and Chiappori, Salanié and Weiss (2015) used the term "separability" for the implementation in Choo and Siow of this sort of distinction between the cardinalities of observable and unobservable agent types. Like the marriage literature on models with a continuum of agents, I will maintain separability in part to use the AH model's greater generality.

Separability in the AH model is defined as follows. Let the valuation of an agent of full type $i \in I$ corresponding to observable and unobservable types $(j, k) \in J \times K$ be

$$v^{i}(\Phi, \Psi) = \pi^{j}(\Phi, \Psi) + \epsilon^{k}_{\Phi,\Psi},$$

where π^{j} is the **valuation function for observable agent type** j over trades (Φ, Ψ) and $\epsilon^{k}_{\Phi,\Psi}$ is the **unobservable valuation component** for unobservable agent type k and the set of trades (Φ, Ψ). Under separability, π^{j} is a function of only the observable agent type j and the sets of trades Φ and Ψ. All aspects of a trade ω are in the data, including the observable agent types of the buyer $b(\omega) \in J$ and seller $s(\omega) \in J$.[5] An agent of full type i also has an unobservable valuation component

[5] Let $\pi^{j}(\Phi, \Psi) = -\infty$ if j is not $b(\omega)$ for any $\omega \in \Phi$ or not $s(\omega)$ for any $\omega \in \Psi$.

over the trades Φ and Ψ, given by $\epsilon^k_{\Phi,\Psi}$. The unobservable valuation components are separate for each set (Φ, Ψ) and so vary for k based on the observable agent types $b(\omega)$ and $s(\omega)$ of the agents on the other side of trades in (Φ, Ψ). However, $\epsilon^k_{\Phi,\Psi}$ does not depend on the unobservable agent types for the counterparties on trades in (Φ, Ψ).

Recall the distribution $\eta(i)$ over agent full agent types $i \in I$. Each full agent type i is also the realization of an observable type j and the realization of an unobservable type k that itself indexes a realization of the vector $\epsilon^k = \left(\epsilon^k_{\Phi,\Psi}\right)_{\Phi \subseteq \Omega, \Psi \subseteq \Psi}$. The vector ϵ^k is of finite length because the set of trades Ω is finite.[6] Therefore, $\eta(i)$ induces a joint distribution (CDF) $F\left(\epsilon^k \mid j\right)$ for each observable agent type j.[7] By this notation, the vector ϵ^k is independently distributed across agents, conditional on the observable types $j \in J$.

3.3 Examples

Example 1. Consider the monogamous, heterosexual marriage setting in Choo and Siow (2006), Galichon and Salanie (2012) and Chiappori, Salanié and Weiss (2015), which uses separability and a continuum of agents. This is an example of one-to-one, two-sided matching.[8] The matching maximum score estimator was already previewed for this example in Section 2. Divide agents into males and females. Each observable agent type j corresponds to a sex (male or female) and other observable demographic characteristics, such as age and race. Age is measured in integer years to have finite support. A trade ω corresponds to a marriage: a male observable type $j_m = b(\omega)$ and a female observable type $j_f = s(\omega)$. The price p_ω of a trade is exchanged between males and females. Define $\pi^j(\Phi, \Psi)$ to be $-\infty$ if any agent is engaged in more than one marriage or married to an agent of the same sex. Therefore, valuations for a male type i or (j, k) from matching with a female observable type $j_f = s(\omega)$ specialize to

$$\pi^j(j_f) + \epsilon^k_{j_f},$$

where $\pi^j(j_f)$ is the valuation function for males with observable demographics in j matching to females with observable demographics j_f and $\epsilon^k_{j_f}$ is the preference of a male of unobservable type k for females with demographic characteristics j_f. A symmetric valuation exists for females of type i. Separability in the marriage setting means that males have preferences over female demographics, not the unobservable type of the female they match with. While not considered in the cited empirical literature, it is straightforward to include aspects other than demographics into a trade ω. For example, a trade ω could specify the number of children or the hours of work of each spouse in a marriage. Then the valuation for a male of type i or (j, k) for trade ω would be

$$\pi^j(\omega) + \epsilon^k_\omega.$$

[6] The notation ϵ^k means the unique realization of the vector of unobservables corresponding to the unobservable agent type k. When k is a random variable, then so is ϵ^k.

[7] One can safely drop the $\epsilon^k_{\Phi,\Psi}$ for a particular j and (Φ, Ψ) when $\pi^j(\Phi, \Psi) = -\infty$ for that j.

[8] The two-sidedness (heterosexual marriage) defines Example 1 and is used in the cited empirical literature. The AH model can also be specialized to one-sided models of marriage: homosexual marriage or a model with both heterosexual and homosexual marriage.

A similar valuation exists for females. Empirical implementation of the more general notion of a trade requires the extra elements to be observable in the data, as in data on labor supply and the number of children for each marriage that occurs in the data.

Example 2. Say an agent is defined to be either a buyer or a seller ex ante, as in the empirical work on the car parts industry later in this paper. Then this is an example of two-sided, many-to-many matching. Define $\pi^j(\Phi, \Psi)$ to be $-\infty$ if an agent whose observable type $j \in J$ corresponds to a buyer conducts trades as a seller, and similarly for a seller type. A trade ω specifies the buyer observable type $b(\omega)$ and the seller observable type $s(\omega)$ in addition to possible other attributes, such as the quantity and quality of goods to deliver (if quantity and quality are specified on a finite grid and observable in the data for actual matches). Under separability, a buyer of full type i or (j, k) then has profits of

$$\pi^j(\Phi) + \epsilon^k_\Phi - \sum_{\omega \in \Phi} p_\omega.$$

As in marriage, the buyer's unobservable valuation component ϵ^k_Φ depends on the trades and hence on the observable types $s(\omega) \in J$ of the seller partners. Similarly, a seller full type i or (j, k) has profits of

$$\pi^j(\Psi) + \epsilon^k_\Psi + \sum_{\omega \in \Psi} p_\omega.$$

Recall that AH prove that a competitive equilibrium exists in this setting without ruling out empirically relevant cases, such as a function $\pi^j(\Phi)$ exhibiting complementarities across multiple trades involving the same agent. Complementarities across multiple trades involving the same agent are vital to the empirical application to the car parts industry.

The AH model requires that an agent's valuation is directly a function of only the trades where that particular agent is a buyer or a seller. The AH model assumes away **externalities**: valuations defined over trades to which the agent does not participate. Competition for trades certainly affects the price vector for trades, p^Ω, although such competition for trades is not a valuation defined over trades to which the agent does not participate. True externalities could be important in applications; for example if buyers are retailers and sellers are wholesalers and buyers compete with each other for retail customers (outside of the matching game) after matching to sellers. Baccara, Imrohoroglu, Wilson and Yariv (2012) use the matching maximum score estimator introduced in this paper to estimate a matching game with externalities.

Example 3. Consider mergers between agents. An agent is not restricted to be a buyer or a seller ex ante. If an agent acquires other agents in equilibrium, it ends up conducting only trades as a buyer although this is not specified ex ante. Likewise, an agent acquired by another agent ends up conducting only a single trade as a seller (if partial acquisitions are not modeled). Therefore, mergers are an example of one-sided matching, also called coalition formation. If desired, one can define $\pi^j(\Phi, \Psi)$ to be $-\infty$ if an agent of type j conducts trades as both a buyer and a seller or if an agent conducts two or more trades as a seller (target). The price p_ω of a trade ω captures the price the buyer (acquirer) pays the seller (target). A trade ω specifies the buyer observable agent type $b(\omega) \in J$ and the seller (target) observable agent type $s(\omega) \in J$. As in Uetake and Watanabe (2012), a trade may also specify other observable attributes, such as the equity split of the post-merger firm or the awarding of board seats to representatives of the acquirer and target. Akkus, Cookson and Hortacsu (forthcoming) use a variant of the matching maximum score estimator with data on the prices of trades p_ω to estimate a matching game of mergers.

As mentioned above, the AH model does not incorporate externalities such as changes in the post-merger competition for customers between firms in the same industry. Therefore, the AH model is a better model of mergers when focusing on, say, across-industry conglomerate mergers.

Example 4. Hatfield et al. (2013) mention the example of trading between dealers of used cars. There is a lively secondhand market in used cars. Dealers may both buy and sell used cars to other dealers. Here a trade ω specifies the observable attributes of the used car in question and buyer and seller observable characteristics $b(\omega), s(\omega) \in J$, including the dealer locations. A buyer might have valuations defined over the location of a seller in order to minimize transportation costs. Dealers might have complex preferences over the set of used cars on their lot. For example, valuations might be higher from ending up with cars of only a certain brand or from having a diverse set of cars. The AH model does not restrict the valuations of dealers over the set of observable trades they undertake. The set of cars that a dealer is endowed with (and possibly does not trade) can be included in the observable agent type $j \in J$.

4 Identification and Estimation

For expositional purposes, I first explore identification and estimation when data on the prices of trades p^Ω are available. This discussion introduces concepts from the literature on single-agent choice and applies them here, although the mathematical results on single-agent choice are not novel. The main purpose of this paper is to study identification and estimation for the case where data on the prices of trades p^Ω are *not* available. The majority of this section discusses the matching maximum score estimator that does not use data on the prices of trades. This estimator was already introduced for the example of marriage in Section 2.

4.1 Single Agent Maximum Score Using Price Data

Define the **choice probability** for observable type j (in equilibrium) to be

$$\bar{A}^j(\Phi, \Psi) = \int_{\epsilon^k} 1 \left[(\Phi, \Psi) \in \arg \max_{\tilde{\Phi} \subseteq \Omega, \tilde{\Psi} \subseteq \Omega} \left(\pi^j \left(\tilde{\Phi}, \tilde{\Psi} \right) + \epsilon^k_{\tilde{\Phi}, \tilde{\Psi}} - \sum_{\omega \in \tilde{\Phi}} p_\omega + \sum_{\omega \in \tilde{\Psi}} p_\omega \right) \right] dF\left(\epsilon^k \mid j\right). \quad (7)$$

This is the same choice probability (or market share equation) from the literature on estimating single agent multinomial choice models (McFadden, 1973). Note that while prices for trades p^Ω are determined in the equilibrium to the AH model, prices are not statistically endogenous in the sense of being statistically dependent with k (or ϵ^k). This is a key implication of the distinction between the finite trades $\omega \in \Omega$ and the infinite full agent types $i \in I$ in the AH model and of the related separability assumption. Because the price vector p^Ω is fixed in a competitive equilibrium, the price vector p^Ω is not listed as an explicit argument to $\bar{A}^j(\Phi, \Psi)$. Note also that the notation $\bar{A}^j(\Phi, \Psi)$ for choice probabilities for observable type $j \in J$ is analogous to the allocation $A^i(\Phi, \Psi)$ for full type $i \in I$. $\bar{A}^j(\Phi, \Psi)$ is observable in the data (with an infinite data set) even when prices p^Ω are not observable and so one cannot use (7) to explicitly compute $\bar{A}^j(\Phi, \Psi)$ given π^j and $F\left(\epsilon^k \mid j\right)$.

For this subsection only, say that prices p^Ω for trades *are* in the data. Say also that the researcher has data on the trades (Φ_i, Ψ_i) for a random sample (i.i.d.) of N agents. The trades involving agent i give us i's observable agent type $j \in J$. Then one can attempt to (point or set) identify π^j for each j up to location and scale normalizations using methods from the single agent multinomial choice literature. To lead up to the matching maximum score estimator, I illustrate the use of single agent methods with price data using multinomial choice maximum score (Manski, 1975; Matzkin, 1993; Fox, 2007).

I will focus on the semiparametric case, where the valuation function

$$\pi^j(\Phi, \Psi) = \pi_\theta(j, \Phi, \Psi)$$

is known up to a finite vector of parameters θ (the "parametric" in "semiparametric") and $F\left(\epsilon^k \mid j\right)$ is not known up to a finite number of parameters for each j (so is nonparametrically specified). If the parameters in π^j vary across j, collect them all in θ. I do not assume that $F\left(\epsilon^k \mid j\right)$ is common across j, so heteroskedasticity is allowed. Allowing heteroskedasticity is a feature of maximum score that is typically ruled out in other semiparametric approaches to discrete choice estimation. The parameterization of $\pi_\theta(j, \Phi, \Psi)$ is for empirical convenience; Matzkin (1993) studies single agent, multinomial choice maximum score estimation when each $\pi^j(\Phi, \Psi)$ is nonparametrically specified and Fox (2010) studies nonparametric identification of aspects of $\pi^j(\Phi, \Psi)$ in a matching model without price data, as discussed below.

To illustrate practical implementation, I further restrict the valuation function to be linear in the parameters θ,

$$\pi_\theta(j, \Phi, \Psi) = X(j, \Phi, \Psi)' \theta,$$

where $X(j, \Phi, \Psi)$ is a vector of observables chosen by the researcher.

As described in Fox (2007), some key conditions for (set) identification and hence (set) consistency of a single agent maximum score estimator are that ϵ^k has **full support** in $\mathbb{R}^{\dim\left(\epsilon^k\right)}$ and that ϵ^k has an **exchangeable distribution** for each j. Let ρ be a permutation of the elements of ϵ^k. An exchangeable distribution satisfies $F\left(\epsilon^k \mid j\right) = F\left(\rho\left(\epsilon^k\right) \mid j\right)$ for all such permutations ρ. Exchangeable distributions allow certain types of equicorrelation across the elements of ϵ^k but rule out some common empirical specifications, such as the random coefficients logit where θ is interpreted as the mean of the random coefficients.

Fix the observable agent type j. Under full support and exchangeability of $F\left(\epsilon^k \mid j\right)$, Goeree et al. (2005) and Fox (2007) show, for the matching notation used here, that a **single agent rank order**

property holds: $\bar{A}^j(\Phi_1,\Psi_1) \geq \bar{A}^j(\Phi_2,\Psi_2)$ if and only if

$$X(j,\Phi_1,\Psi_1)'\theta - \sum_{\omega\in\Phi_1} p_\omega + \sum_{\omega\in\Psi_1} p_\omega \geq X(j,\Phi_2,\Psi_2)'\theta - \sum_{\omega\in\Phi_2} p_\omega + \sum_{\omega\in\Psi_2} p_\omega.$$

Roughly speaking, one can interpret $X(j,\Phi,\Psi)'\theta$ as the mean valuation of the observable type j and the rank order property says that choices with higher mean valuation plus prices are made more often. The single agent rank order property is a statement about the properties of the exchangeable distribution $F\left(\epsilon^k \mid j\right)$ and how the unobservables enter the choice model.

Assuming that the elements of $X(j,\Phi,\Psi)$ are linearly independent, Manski (1975) and Fox (2007) show that θ is point identified if one element of the vector $X(j,\Phi,\Psi)$ has **full support** (equal to $\mathbb{R}$) conditional on the other elements of $X(j,\Phi,\Psi)$, on the price vector p^Ω (which are not random variables in a competitive equilibrium), and also on the vectors $X(j,\Phi,\Psi)$ for other sets of trades (Φ,Ψ). In the AH model, the set of trades Ω is finite so that there is no story in the model where any element of $X(j,\Phi,\Psi)$ could have support on an interval in $\mathbb{R}$. Formally speaking, if Ω is indeed finite than θ will be set identified. Bajari, Fox and Ryan (2008) estimate a multinomial choice maximum score model and conduct inference allowing for set identification using the method of Romano and Shaikh (2008). The empirical marriage matching model of Dupuy and Galichon (2014) allows for continuous observable characteristics and hence θ could be point identified under that model.

Say agent i of observable type j_i undertakes the trades (Φ_i,Ψ_i) in the data. The chosen sets of trades (Φ_i,Ψ_i) represent the dependent variable in the multinomial choice model. Let B be a set of sets of possible trades (Φ,Ψ). The single agent, pairwise maximum score objective function in Fox (2007), which here uses data on prices p^Ω, is then

$$\sum_{i=1}^{N} \sum_{(\Phi,\Psi)\in B} 1\left[X(j_i,\Phi_i,\Psi_i)'\theta - \sum_{\omega\in\Phi_i} p_\omega + \sum_{\omega\in\Psi_i} p_\omega \geq X(j_i,\Phi,\Psi)'\theta - \sum_{\omega\in\Phi} p_\omega + \sum_{\omega\in\Psi} p_\omega\right]. \tag{8}$$

The estimator $\hat{\theta}_N$ is any vector that maximizes the objective function. The maximum score objective function using price data is a step function and so a continuum of parameter values will maximize it; any rule to pick one (such as letting a numerical optimization routine decide) leads to a consistent estimator under point identification. Manski (1975) and Fox (2007) demonstrate the consistency of the single agent estimator as $N\to\infty$. For the point identified case, Kim and Pollard (1990) present the asymptotic distribution and Delgado, Rodriguez-Poo and Wolf (2001) discuss subsampling for straightforward inference. As just mentioned, an inference method such as Romano and Shaikh (2008) is applicable to the set identified case. Akkus, Cookson and Hortacsu (forthcoming) and, later, Fox and Bajari (2013, Section VI.B, Appendix C) estimate matching models using single agent maximum score estimators using price data and perform Monte Carlo studies. They do not formally state assumptions on the competitive equilibrium for an underlying matching model, such as the AH model, in order to motivate the consistency of the single agent maximum score estimator using primitive assumptions on $F\left(\epsilon^k \mid j\right)$.

The single agent maximum score estimator using data on prices p^Ω has many of the same attractive computational properties as the matching maximum score estimator previewed in Section 2 for marriage without data on prices p^Ω. Most notably, there is no numerical integration over the typically very long vector ϵ^k and no nested algorithm to compute a competitive equilibrium. Also, there is no data curse of dimensionality from estimating $F\left(\epsilon^k \mid j\right)$ using a sieve approximation (Chen, 2007) or from estimating choice probabilities $\bar{A}^j(\Phi, \Psi)$ in a nonparametric first stage.[9] The maximum score estimator using data on prices is consistent if B is not the set of all sets of trades (Φ, Ψ), which is convenient as that set, $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$, is typically quite large in the AH model and other matching models (Fox, 2007). Itemizing all elements of $\mathcal{P}(\Omega) \times \mathcal{P}(\Omega)$ is computationally prohibitive in many applications.

Discrete choice models require location and scale normalizations. The scale normalization for single agent choice is that the parameters are in monetary units, as the price vector p^Ω is observed. This scale normalization will change when prices p^Ω are not observed below. The location normalization depends on whether the outside option of making no trades is included in the set B. If not, no constant term should be included in $X(j, \Phi, \Psi)$ as the constant would be differenced out in the maximum score inequalities. If the outside option of making no trades is both observed in the data (agents with no trades are measured) and is normalized to have a profit of 0, then a constant term (possibly a separate one for different classes of observable agent types $j \in J$, such as males and females in marriage) can be included in $X(j, \Phi, \Psi)$.

4.2 Nonparametric Identification with Unobservable Prices

Nonparametric identification in a transferable utility matching game seeks to learn aspects of $\pi^j(\Phi, \Psi)$ without specifying $\pi^j(\Phi, \Psi)$ up to a finite vector of parameters θ. Nonparametric identification when prices of trades p^Ω are unobservable in the data differs greatly from the analysis of nonparametric identification of $\pi^j(\Phi, \Psi)$ in single agent choice under maximum score like assumptions in Matzkin (1993), at least under the AH model and the separability assumption.

Loosely speaking, the classic paper of Becker (1973) initiated the study of nonparametric identification of aspects of the valuation functions using data on who matches with whom and where prices are present in the model but are not in the data. Becker studied marriage, or one-to-one, two-sided matching. In his model, each male and female has a scalar observable type on $\mathbb{R}$, say j_m for males and j_f for females. The scalar type could be schooling for males and schooling for females. There is no unobservable type k. The observed, scalar types enter a **joint valuation function** $g(j_m, j_f)$ that, under transferable utility, is equal to $\pi^{j_m}(j_f) + \pi^{j_f}(j_m)$, the sum of the valuations of a male and a female. Becker restricts attention to the case where either j_m and j_f are substitutes or they are complements. The scalar types of two matched agents are **complements** when

$$\frac{\partial^2 g(j_m, j_f)}{\partial j_m \partial j_f} > 0$$

globally and **substitutes** when the cross-partial derivative is negative. In a special case of the efficiency result for the AH model mentioned earlier, Becker shows that complementarities lead to

[9]The distribution $F\left(\epsilon^k \mid j\right)$ is not estimated, as is standard in maximum score. Knowledge of such a distribution is needed to compute marginal effects and elasticities. After θ is estimated using single agent maximum score, a sieve maximum likelihood approach (under perhaps stronger assumptions such as no heteroskedasticity) could be used to estimate the distribution of ϵ^k (Chen, Tamer and Torgovitsky, 2011).

positive assortative matching: high j_m males marry high j_f females. Substitutes lead to **negative assortative matching**. Therefore, if a researcher observes positive assortative matching in the data, then one might conclude that the cross-partial derivative of the joint valuation function is positive. One can identify only the sign of the cross-partial derivative in Becker's setup.

Fox (2010) studies nonparametric identification of joint valuation functions in a two-sided, many-to-many matching game using a rank order property for matching games with a finite number of agents. He uses data on who matches with whom but not prices. He extends the intuition in Becker as well as some results in Matzkin (1993) on single agent choice to the many-to-many matching case where each agent (or group of agents) has a vector of observable characteristics. For example, one can identify the ratio of complementarities (cross partials) on two matched agent characteristics to the complementarities on two other matched agent characteristics. One identifies the numerical value of the ratio, not just the sign of the ratio. Identification results vary somewhat depending on whether data on unmatched agents are used.

As Fox (2010) shows for two-sided, many-to-many matching, results on the nonparametric identification of aspects of $\pi^j(\Phi, \Psi)$ in the AH model using data on trades and not prices do not always equal results from using a single agent analysis with data on the prices of trades p^Ω (Matzkin, 1993). However, the current paper is on tractable semiparametric estimation of θ in the general setting of the AH model, not nonparametric identification of aspects of $\pi^j(\Phi, \Psi)$ in a more specific setting. Ideally, the elements of $X(j, \Phi, \Psi)$ should be chosen based on intuition arising from formal results on nonparametric identification of $\pi^j(\Phi, \Psi)$ in the special case of the AH model being estimated. In what follows, I formally establish only set identification of θ but am motivated by models where the elements of $X(j, \Phi, \Psi)$ are specified in such a way that the coming matching maximum score inequalities do not always difference out an element of the vector θ. Without data on prices, a scale normalization on the vector θ is needed. I assume that one element of θ is either $+1$ or -1. The sign of θ will typically be identifiable from the data.

4.3 Matching Maximum Score

To use a maximum score estimator without data on the prices of trades p^Ω, in this subsection I use intuition to define matching maximum score inequalities and the matching maximum score objective function. In the next subsection, I prove a rank order property for matching without data on prices that a further result uses to prove that the matching maximum score inequalities lead to set identification of the parameter vector θ.

For intuition, a convenient property of the AH model is that the allocation portion A of any competitive equilibrium $\left(A, p^\Omega\right)$ is efficient in the sense of maximizing the social welfare function (6). Therefore, any other allocation should weakly lower social welfare. First consider the full version of the AH model where the same pair of two agents can undertake two different trades. In this case, we can base an inequality around two specific trades, ω_1 and ω_2. The multi-set (allowing duplicates in the set) of the buyer and seller observable types for trade ω_1, $b(\omega_1), s(\omega_1) \in J$, must equal the multi-set of the buyer and seller observable types for trade ω_2. The buyer observable type $b(\omega_1)$ on trade ω_1 could be either the buyer or the seller observable type on trade ω_2, although a particular maximum score inequality fixes the role of $b(\omega_1)$ on trade ω_2

The deviation from trade ω_1 to trade ω_2 for the two observable types $b(\omega_1), s(\omega_1) \in J$ is feasible because there is both a buyer and a seller for each trade under both circumstances. On the left side of the inequality, agent $b(\omega_1)$ conducts the total trades $\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}\right)$ and agent $s(\omega_1)$ conducts the total trades $\left(\Phi_{s(\omega_1)}, \Psi_{s(\omega_1)}\right)$. Further, let $\left(\bar{\Phi}_{s(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right)$ and $\left(\bar{\Phi}_{s(\omega_1)}, \bar{\Psi}_{s(\omega_1)}\right)$ be the respective trades when the agents $b(\omega_1), s(\omega_1) \in J$ switch from trade ω_1 to ω_2. Then a **matching maximum score inequality** based on trades ω_1 and ω_2 (and on $\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}\right)$ and $\left(\Phi_{s(\omega_1)}, \Psi_{s(\omega_1)}\right)$) is

$$X\left(b(\omega_1), \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}\right)' \theta + X\left(s(\omega_1), \Phi_{s(\omega_1)}, \Psi_{s(\omega_1)}\right)' \theta \geq X\left(b(\omega_1), \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right)' \theta + X\left(s(\omega_1), \bar{\Phi}_{s(\omega_1)}, \bar{\Psi}_{s(\omega_1)}\right)' \theta. \tag{9}$$

The intuition behind the inequality is that the social welfare for trade ω_1 must be greater than the social welfare when the agent observable types instead engage in trade ω_2. This motivation is only intuition as the inequality drops the unobservable agent types k (the unobservables in each ϵ^k) and so we must prove a rank order property to show that a maximum score estimator based on this inequality will set identify the true θ.

In some examples of applying the AH model, the inequality (9) will not be informative. Returning to Example 1, consider one-to-one, two-sided matching (marriage) where trades ω encode *only* the observable agent types of the buyer and the seller. As trades encode no other features than observable types $j \in J$, a male observable type conducting a marriage trade $\omega_1 \in \Omega$ with a female observable type cannot instead conduct a distinct marriage trade $\omega_2 \neq \omega_1$ with that same female observable type. Recall that Section 2 presents a matching maximum score inequality for marriage, (1). We now introduce notation for a matching maximum score inequality that generalizes the inequality (1) for the marriage example as well as the inequality (9) just introduced.

Let the more general matching maximum score inequality be indexed by g out of some finite set G of possible inequalities. The set G is finite as the set of trades Ω is finite in the AH model. The set of trades Ω might be infinite in some other matching model like Dupuy and Galichon (2014); this is not a challenge for maximum score.

An inequality g will focus on the two trades ω_1 and ω_2 in the multi-set (allowing duplicates) $\Omega_g = \{\omega_1, \omega_2\}$ on the inequality's left side and the two other trades $\bar{\Omega}_g = \{\omega_3, \omega_4\}$ on the inequality's right side. The trades can include the option of not making a trade in order to explore agents dropping or adding trades. The set of observable types of agents should be the same for the left and right sides: the multi-set $H_g = \{b(\omega_1), s(\omega_1), b(w_2), s(\omega_2)\}$ should equal the multi-set $\bar{H}_g = \{b(\omega_3), s(\omega_3), b(w_4), s(\omega_4)\}$. Further, there is some unique mapping between the agents in H_g and $\bar{H}_g$ in the case of identical observable agent types. For each observable type $j \in H_g$, let (Φ_j, Ψ_j) be $j \in H_g$'s total trades on the left side of the inequality; the corresponding trade $\omega \in \Omega_g$ where j is a buyer or seller must be in (Φ_j, Ψ_j). Likewise, each $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)$ is $j \in H_g$'s total trades on the right side of the inequality, where the corresponding trade $\omega \in \bar{\Omega}_g$ where j is a buyer or seller must be in $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)$ and $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)$ must be equal to (Φ_j, Ψ_j) for each $j \in H_g$ except for the trade $\omega \in \bar{\Omega}_g$ replacing the corresponding trade $\omega \in \Omega_g$.[10] Given this notation, the **matching maximum score inequality** g based on the trades ω_1–ω_4 and the corresponding sets (Φ_j, Ψ_j) and $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)$ for $j \in H_g$ is defined to

[10]The trade sets (Φ_j, Ψ_j) and $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)$ should be subscripted by the inequality index g; this is dropped for conciseness.

be

$$X\left(b\left(\omega_{1}\right),\Phi_{b(\omega_{1})},\Psi_{b(\omega_{1})}\right)'\theta+X\left(s\left(\omega_{1}\right),\Phi_{s(\omega_{1})},\Psi_{s(\omega_{1})}\right)'\theta+X\left(b\left(\omega_{2}\right),\Phi_{b(\omega_{2})},\Psi_{b(\omega_{2})}\right)'\theta+X\left(s\left(\omega_{2}\right),\Phi_{s(\omega_{2})},\Psi_{s(\omega_{2})}\right)'\theta\geq$$
$$X\left(b\left(\omega_{1}\right),\bar{\Phi}_{b(\omega_{1})},\bar{\Psi}_{b(\omega_{1})}\right)'\theta+X\left(s\left(\omega_{1}\right),\bar{\Phi}_{s(\omega_{1})},\bar{\Psi}_{s(\omega_{1})}\right)'\theta+X\left(b\left(\omega_{2}\right),\bar{\Phi}_{b(\omega_{2})},\bar{\Psi}_{b(\omega_{2})}\right)'\theta+X\left(s\left(\omega_{2}\right),\bar{\Phi}_{s(\omega_{2})},\bar{\Psi}_{s(\omega_{2})}\right)'\theta. \tag{10}$$

The inequality states that the sum of the valuation functions from the two trades in $\Omega_g = \{\omega_1, \omega_2\}$ is greater than the sum of the valuation functions from the two trades in $\bar{\Omega}_g = \{\omega_3, \omega_4\}$.

If $\omega_1 = \omega_2$ and $\omega_3 = \omega_4$, this definition encompasses (9); the inequality (9) would be the appropriate (10) divided by 2 on both sides. For the marriage example in Section 2, let

$$\tilde{X}(j_m, j_f) = X\left(b\left(\omega_1\right), \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}\right) + X\left(s\left(\omega_1\right), \Phi_{s(\omega_1)}, \Psi_{s(\omega_1)}\right)$$

for trade ω_1 where $b(\omega_1) = j_m$ and $s(\omega_1) = j_f$. Hence the marriage matching maximum score inequality (1) is a special case of (10).

The researcher chooses the set G of possible matching maximum score inequalities to use in estimation. A *possible* inequality $g \in G$ becomes an *actual* inequality in estimation whenever the configuration of observable agent types and sets of trades on the left side of the inequality (10) is sampled in the data.

The researcher has a lot of freedom to choose the set G of possible matching maximum score inequalities. However, for set identification, discussed below, there are some conditions needed on the (asymptotic) choice of matching maximum score inequalities. Each inequality $g_1 \in G$ has a $\geq$ sign, as in (10). Let $g_2(g_1)$ be the reverse-direction inequality: the same inequality with a $<$ sign, which is equivalent to switching the right and left sides of g_1. The condition needed is that whenever the inequality $g_1 \in G$, then $g_2(g_1) \in G$. In other words, g_1 is included whenever the multi-sets of trades Ω_{g_1} (and $(\Phi_j, \Psi_j)_{j \in H_{g_1}}$) are sampled in the data and likewise the reverse-direction inequality $g_2(g_1)$ is included whenever the multi-sets of trades $\bar{\Omega}_{g_1} = \Omega_{g_2}$ (and $\left(\bar{\Phi}_j, \bar{\Psi}_j\right)_{j \in \bar{H}_{g_1}}$) are sampled in the data.[11]

The matching maximum score inequality g in (10) can be notationally simplified. The parameter vector θ multiplies all four X vectors in the inequality. Therefore we can collect terms by defining the vector

$$Z_g = \sum_{\omega \in \Omega_g} \left(X\left(b\left(\omega\right), \Phi_{b(\omega)}, \Psi_{b(\omega)}\right) + X\left(s\left(\omega\right), \Phi_{s(\omega)}, \Psi_{s(\omega)}\right)\right) - \sum_{\omega \in \bar{\Omega}_g} \left(X\left(b\left(\omega\right), \bar{\Phi}_{b(\omega)}, \bar{\Psi}_{b(\omega)}\right) + X\left(s\left(\omega\right), \bar{\Phi}_{s(\omega)}, \bar{\Psi}_{s(\omega)}\right)\right).$$

Then the matching maximum score inequality g in (10) can be written as $Z_g'\theta \geq 0$.

The matching maximum score objective function for a sample of data on the trades (Φ_i, Ψ_i) of

[11] More formally, to rule out two otherwise mutually exclusive inequalities being true for the same parameter value θ and neither being true, assign one of each pair $g_1 \in G$ and $g_2(g_1) \in G$ a weak inequality $\geq$ and the other one a strict inequality $>$.

$i = 1, \ldots, N$ agents, but not the prices of trades p^{Ω}, is

$$\sum_{g \in G_N} 1 \left[Z_g' \theta \geq 0 \right], \tag{11}$$

where G_N are the inequalities to use for this sample. G_N may be a multi-set as the same inequality could appear multiple times if agents of the same observable type $j \in J$ are observed. Computationally, this linear-in-parameters form is the same form as the inequalities in single agent maximum score and maximum rank correlation estimators (Manski, 1975, 1985; Han, 1987; Fox, 2007). As stated before, the matching maximum score estimator avoids nonparametric estimates of choice probabilities and distributions of unobservables, numerical integration, and algorithms to compute competitive equilibria. The set G of possible inequalities can be chosen for computational convenience. Therefore, the matching maximum score estimator is practical when certain alternatives are not.

The matching maximum score objective function (11) can be rewritten in a way that facilitates calculating its expectation and its probability limit under i.i.d. sampling of the trades of agents $i \in I$. With an appropriate normalizing constant, the matching maximum score objective function is also, for $N \geq 4$,

$$\binom{N}{4}^{-1} \sum_{i_1=1}^{N-3} \sum_{i_2=i_1+1}^{N-2} \sum_{i_3=i_2+1}^{N-1} \sum_{i_4=i_3+1}^{N} \sum_{g \in G} 1 \left[\{(\Phi_i, \Psi_i)\}_{i=i_1,i_2,i_3,i_4} = \left\{ \left(\Phi_j^g, \Psi_j^g \right) \right\}_{j \in H_g} \right] 1 \left[Z_g' \theta \geq 0 \right]. \tag{12}$$

Here the outer four summations form all sets of four agents i in the sample. The inner sum is over all matching maximum score inequalities $g \in G$. For each inequality g, the objective function checks with the multi-set of trades of the four agents is equal to the multi-set of trades on the left side of the inequality g. A trade ω contains the observable types of the buyer and seller, and so the objective function also checks if the observable agent types match the inequality g. If the inequality checks out, then the corresponding matching maximum score inequality is included in the maximum score objective function as the indicator

$$1 \left[\{(\Phi_i, \Psi_i)\}_{i=i_1,i_2,i_3,i_4} = \left\{ \left(\Phi_j^g, \Psi_j^g \right) \right\}_{j \in H_g} \right] \tag{13}$$

equals 1 and so the inequality itself, $Z_g' \theta \geq 0$, enters the objective function. The model's dependent variable, the trades (Φ_i, Ψ_i) that are undertaken by the N agents in the data, enters the matching maximum score objective function (12) through (13).

4.4 Rank Order Property and Set Identification

The parameter vector θ is set identified using data on trades and not the prices of trades p^{Ω} when a rank order property holds. The following proposition defines the **rank order property** for matching without data on the prices of trades and states that it holds. Recall that $\bar{A}^j(\Phi, \Psi)$ is the fraction or choice probability of agents of observable type $j \in J$ that conduct the trades Φ and Ψ. Let

$$\bar{A}^j(\Phi, \Psi \mid \Phi_1, \Psi_1, \Phi_2, \Psi_2) = \frac{\bar{A}^j(\Phi, \Psi)}{\bar{A}^j(\Phi_1, \Psi_1) + \bar{A}^j(\Phi_2, \Psi_2)}$$

be the probability that observable type $j \in J$ picks the trades (Φ, Ψ) conditional on the event that j picks either the trades (Φ_1, Ψ_1) or the trades (Φ_2, Ψ_2). Assume that $F\left(\epsilon^k \mid j\right)$ has bounded, continuous derivatives, in addition to having full support on $\mathbb{R}^{\dim\left(\epsilon^k\right)}$ and being exchangeable, as before.

Proposition 1. *The matching maximum score inequality* (12) *holds if and only if the following inequality holds:*

$$\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right) \cdot \bar{A}^{b(\omega_2)}\left(\Phi_{b(\omega_2)}, \Psi_{b(\omega_2)} \mid \Phi_{b(\omega_2)}, \Psi_{b(\omega_2)}, \bar{\Phi}_{b(\omega_2)}, \bar{\Psi}_{b(\omega_2)}\right) \geq \bar{A}^{b(\omega_1)}\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right) \cdot \bar{A}^{b(\omega_2)}\left(\bar{\Phi}_{b(\omega_2)}, \bar{\Psi}_{b(\omega_2)} \mid \Phi_{b(\omega_2)}, \Psi_{b(\omega_2)}, \bar{\Phi}_{b(\omega_2)}, \bar{\Psi}_{b(\omega_2)}\right). \tag{14}$$

The probability statement involves only buyer probabilities as buyer probabilities are related to seller probabilities by feasibility, (5).[12] It is important that the same two observable agent types' choice probabilities are on the left and right sides of the probability statement (14). Recall that the buyer on, for example, trade $\omega_1 \in \Omega_g$ might be a seller on trade $\omega_3 \in \bar{\Omega}_g$.

In words, the rank order property for matching without price data states that the conditional probability of observing the configuration of trades on the left side of (10) is greater than the conditional probability of observing the configuration of trades on the right side of (10) whenever the sum of valuations involving observable types j and trades ω on the left side of (10) exceed those on the right side of (10). The rank order property allows an estimator based on maximizing matching maximum score inequalities involving only measured observable agent types j and trades ω to be (set) consistent in the presence of unobservable types k, which index the vectors ϵ^k.

The matching maximum score inequality for Example 1, marriage without price data, is (1). Applied to marriage, Proposition 1 states that the inequality (1) holds for observable types, in a slight adjustment of notation, of the form $j_m^{i_i} = j_m^1$ if and only if

$$A^{j_m^1}\left(j_f^1 \mid j_f^1, j_f^2\right) \cdot A^{j_m^2}\left(j_f^2 \mid j_f^1, j_f^2\right) \geq A^{j_m^1}\left(j_f^2 \mid j_f^1, j_f^2\right) \cdot A^{j_m^2}\left(j_f^1 \mid j_f^1, j_f^2\right)$$

where $A^{j_m}\left(j_f^1 \mid j_f^1, j_f^2\right)$ is the probability of observable male type j_m picking observable female type j_f^1 conditional on picking either female type j_f^1 or female type j_f^2. The simple matching maximum score inequality for two trades without price data is (9). Using that inequality's notation, Proposition 1 plus algebraic simplification states that the inequality (9) holds if and only if

$$\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right) \geq \bar{A}^{b(\omega_1)}\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right),$$

as there is only one trade on each side of the inequality (9).

The proof of Proposition 1 is in an appendix. A full proof of the proposition uses the assumption from single-agent maximum score that $F\left(\epsilon^k \mid j\right)$ has full support and is an exchangeable distribution for all j to apply the single agent rank order property for multinomial choice, mentioned above (Manski, 1975; Fox, 2007). In addition to single agent maximum score results, the full proof of the theorem also uses properties that only hold in competitive equilibrium; thus the proof uses matching

[12] The trades of sellers in the inequality g other than ω_1–ω_4 contribute to the sampling of inequalities in the data. This does not affect the statement of the rank order property, as the proof indicates.

theory in addition to manipulating the integrals in the definition of a choice probability (7). The survey of Graham (2011, Theorem 4.1) was the first to prove that the rank order property holds in a semiparametric model of marriage. The intellectual history of this is tangled. Previous circulating working paper versions of the current paper, starting in 2004, introduced the matching maximum score estimator. Later working paper versions pointed out that the rank order property was implied by the closed form allocation (matching) probabilities in the parametric Choo and Siow (2006) model of marriage, which uses the type I extreme value (logit) distribution for each element of the vector ϵ^k. However, the proof idea for the rank order property for the semiparametric case originates in Graham, who extends the results on one-to-one, two-sided matching with a continuum of agents to the semiparametric setting, namely the assumption that $F\left(\epsilon^k \mid j\right)$ is an exchangeable distribution for all j. Proposition 1 extends Graham's result from marriage to the full generality of matching with trades in the AH model. To clarify the intellectual contribution, the proof of Proposition 1 in the appendix cites Graham's result for marriage and shows how the inequalities in Graham can be modified to derive the probability statement (14) in the AH model.[13]

Using the rank order property in Proposition 1, we can prove that the model is **set identified**, meaning that the set of maximizers of the expectation of the maximum score objective function contains the true parameter. The expectation is equal to the probability limit based on i.i.d. sampling of agents i and their trades (Φ_i, Ψ_i).

Theorem 1. *The set of maximizers θ of the expectation of the maximum score objective function* (1) *contains the true parameter vector.*

Say one further imposes that an element of the vector Z_g has support on $\mathbb{R}$ conditional on other elements of Z_g, contradicting the finite set of trades Ω in the AH model, but not the marriage model of Dupuy and Galichon (2014). The proof of the (point) consistency of maximum score would then verify that set identification reduces to point identification and that other conditions in a general consistency theorem for extremum estimators are satisfied (Newey and McFadden, 1994).[14] Fox (2007) provides such a proof for single-agent, multinomial choice maximum score and all the steps go through for the matching case as well.

4.5 Inference in Matching Maximum Score

As mentioned previously, set inference can use a method such as the subsampling approach of Romano and Shaikh (2008), which was used in single agent maximum score under set identification by Bajari, Fox and Ryan (2008). An input into a procedure such as Romano and Shaikh is the rate of convergence of the objective function.

At least under point identification, the rate of convergence will depend on how quickly the number of inequalities increase with N. In (12), the number of inequalities is proportional to $\begin{pmatrix} N \\ 4 \end{pmatrix}$, which

[13]Graham's proof works by inverting choice probabilities. See the erratum Graham (2013). Note that Graham (2011, Theorem 4.1) is stated for the marriage equivalent of independent and identical $\epsilon^k_{\Phi,\Psi}$ conditional on j instead of an exchangeable $F\left(\epsilon^k \mid j\right)$. The first two steps of Graham's proof reproduce Manski (1975) and Fox (2007), so the assumption of an exchangeable $F\left(\epsilon^k \mid j\right)$ can be used with little change, as in Fox (2007). Graham (2011, Theorem 4.1) allows heteroskedasticity as well.

[14]Also, the elements of Z_g should be linearly independent.

is on the order of N^4. The binomial coefficient $\begin{pmatrix} N \\ 4 \end{pmatrix}$ comes from sampling all sets of four unique agents.

The matching maximum score objective function for marriage with data from a single market is (2). In this example, the number of inequalities is proportional to $\begin{pmatrix} \tilde{N} \\ 2 \end{pmatrix}$, which is on the order of $\tilde{N}^2$, where $\tilde{N}$ is the number of matches (marriages). This is smaller than $\begin{pmatrix} N \\ 4 \end{pmatrix}$ because, in marriage, when one samples a husband one almost always samples the corresponding wife, so it is less natural to view the husband and wife as independent draws from the population of all agents.

In both the objective functions (12) and (2), one can recognize that the objective function has the same double (or more) summations as the maximum rank correlation objective function of Han (1987). Consider the marriage case (2). Sherman (1993, 1994) shows that the maximum rank correlation estimator is $\sqrt{\tilde{N}}$ consistent (under point identification) and asymptotically normal. The limiting variance might be difficult to work with. Under point identification, Subbotin (2007) shows that the bootstrap is valid for inference. Under set identification, the code of Santiago and Fox (2009) in part implements the method of Romano and Shaikh (2008) to construct valid 95% confidence intervals using the rate of convergence $\sqrt{\tilde{N}}$ as an input.

4.6 Multiple Markets

Let there now be data on D markets, indexed by d. Then the matching maximum score objective function (11) has an extra summation over markets,

$$\sum_{d=1}^{D} \sum_{g \in G^d_{N_d}} 1\left[Z'_{g,d}\theta \geq 0 \right], \tag{15}$$

where now the set of inequalities to include $G^d_{N_d}$ is specific to market d.

Let us maintain the AH model: each market is truly a continuum of agents and we merely have data on a subset of N_d agents from market d. There are two asymptotic arguments. The first asymptotic argument fixed the number of markets D and increases the number of agents with recorded data N_d for each d by some common amount $\bar{N}$ such that $N_d = \nu_d \cdot \bar{N}$, for the fixed-with-$\bar{N}$ and market-specific proportionality constants ν_d. Consider the marriage matching maximum score objective function with multiple markets (3). This objective function has the "double summation" form for each market D separately, and the estimator will have the maximum rank correlation asymptotics in some notion of the number of agents in each market $\bar{N}$, meaning the rate of convergence will be $\sqrt{\tilde{N}}$.[15]

Now consider fixing the number of agents with measured data N_d in each market. Each true matching market is still a continuum. A second asymptotic argument makes the number D of markets increase to infinity. Compared to the expectation of the single-market matching maximum score objective function in the proof of Theorem 1, the expectation of (15) involves an outer expectation over

[15] When proving that the true parameter maximizes the probability limit of the objective function, as in Theorem 1, one applies the rank order property in Proposition 1 to inequalities from each market separately.

the distribution of full agent types $\eta_d(i)$ for market d and hence the resulting competitive equilibrium. This extra outer expectation over $\eta_d(i)$ does not change the conclusion of Theorem 1: the true parameter value is a global maximizer of the expectation of the objective function.

For the matching maximum score objective function for marriage in (2), the number of terms in each double summation does not increase with D and so the estimator will have the maximum score asymptotics in Kim and Pollard (1990). In the case of point identification, the estimator will converge as the rate of $D^{1/3}$. Delgado, Rodriguez-Poo and Wolf (2001) show that subsampling is valid for inference. Under set identification, the code of Santiago and Fox (2009) in part implements the method of Romano and Shaikh (2008) to construct valid 95% confidence intervals using the rate of convergence $D^{1/3}$ as an input.[16]

Maximum score estimators allow heteroskedasticity. In matching, even if observable type $j \in J$ has the same meaning across markets, the distribution of unobservables for j need not be the same across markets. Notate the distribution of unobservables as $F\left(\epsilon^k \mid j, d\right)$ for observable type $j \in J$ in market $d \in D$.

4.7 Truly Finite Markets

The Azevedo and Hatfield (2015, Section 6) model with a continuum of agents has useful properties. A competitive equilibrium always exists and is generically unique without strong restrictions on the valuations of agents. The equivalent to the Azevedo and Hatfield (2015, Section 6) model for a market with a **truly finite** number of agents is the model of Hatfield, Kominers, Nichifor, Ostrovsky and Westkamp (2013). I write "truly finite" to emphasize that the issue with finiteness is not about the data; the real equilibrium model in the data generating process has a finite number of agents. A competitive equilibrium to the model of Hatfield et al. (2013) exists if the indirect utility versions of the direct utility valuation functions $v^i(\Phi, \Psi)$ are **submodular**, what the authors call a **substitutes** condition. In one-to-one matching agents can make only one trade and so valuations are a function of a single trade and so trivially submodular. In one-to-many and more complex forms of matching, submodularity of the indirect valuation functions rules out many forms of complementarities across multiple trades involving the same agent, such as the complementarities critical to the empirical study of the car parts industry below. While any competitive equilibrium is efficient and hence the corresponding allocation is unique if the maximizer of the social welfare function is unique, Hatfield et al. demonstrate by example that a failure of the submodularity condition can lead to non-existence of a competitive equilibrium.

As individual matching markets grow large, the finite-agent model of Hatfield et al. (2013) will presumably converge to the continuum model of Azevedo and Hatfield (2015, Section 6), if the separability assumption is maintained. Therefore, a practical suggestion is to treat the matching maximum score estimator as a large market approximation. Instead, Fox (2010) investigates nonparametric iden-

[16] For the example of single agent binary choice, sufficient conditions for set identification in the maximum score model of Manski (1975) allow for heteroskedasticity (here $F\left(\epsilon^k \mid j\right)$ varies with j) while known sufficient conditions for set identification in the maximum rank correlation model of Han (1987), applied to binary choice, require homoskedasticity (here $F\left(\epsilon^k \mid j\right)$ does not vary with j). In matching, the "maximum score" and "maximum rank correlation" asymptotic arguments both allow for heteroskedasticity, based on the rank order property in Proposition 1. The proof of Proposition 1 relies on the properties of competitive equilibrium and so the theorem is not an analog to the conditions for set identification for binary choice in Han (1987).

tification of aspects of the many-to-many joint valuation function in a truly finite market setting by imposing a rank order property for matching without data on prices for the truly finite market setting. While Fox does not discuss estimation or derive a rank order property using primitive conditions on the distribution of unobservables, the rank order property in Fox would lead to the consistency of the matching maximum score estimator in a truly finite market setting. That rank order property can also be applied to the case where the market is truly finite but not all agents are recorded in the data.

5 Car Parts Industry

I now present an empirical application about the matching of assemblers to car parts suppliers 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 trade $\omega \in \Omega$ encodes an individual car part that is needed for a named car model as well as buyer and seller observable types. Each car model itself has a brand. For the Chevrolet Impala, Chevrolet is the brand and Impala is the model. There are multiple consummated trades ω for the Chevrolet Impala because each model uses multiple parts. Finally, each brand is owned by an assembler, in Chevrolet's case General Motors. General Motors is the buyer observable type $b(\omega) \in J$ on all the trades for car parts used on the Chevrolet Impala. The seller observable type $s(\omega) \in J$ on each trade is a particular car parts supplier, like Bosch. Therefore, each named firm represents a separate observable type in the AH model.[17] In the AH model, there is one trade $\omega \in \Omega$ for the windshield on the Chevrolet Impala for each supplier $s(\omega)$ that supplies at least one windshield for any of the models in the data and so could (according to the model) counterfactually supply the windshield for the Impala.

I refer to Example 2 and model the car parts industry as an explicitly two-sided market, where each supplier conducts trades only as a seller and each assembler conducts trades only as a buyer. From an automotive engineering perspective, an assembler needs a specific set of car parts to make a particular model. For example, each car model needs a single windshield.[18] For an assembler i, define $v^i(\Phi) = -\infty$ for any set of trades Φ that do not contain exactly one trade for every car part opportunity in the data. On the right side of a matching maximum score inequality (10), I drop inequalities where $v^i\left(\bar{\Phi}\right) = -\infty$. Therefore, a counterfactual trade ω_3 corresponds to, for example, General Motors using a different supplier for the windshield for the Chevrolet Impala, not General

[17] The economic questions considered here focus on supplier and assembler specialization and so I need to allow each firm to be its own observable type $j \in J$ to properly measure specialization. The AH model uses a finite number of trades $\omega \in \Omega$ and a continuum of full agent types $i \in I$. In my empirical version of the AH model, the number of observable types $j \in J$ is also finite. Here, the fiction mapping the continuum AH model to the finite data is that there is a continuum of firms of the General Motors observable type but only one such firm is sampled in the data.

[18] The data do not report back up or secondary suppliers for a part on a particular car model.

Motors installing two different windshields on the Chevrolet Impala or replacing the windshield with a tire. By the market definition discussed below, only suppliers that make at least one windshield in the data can be the seller observable type on the counterfactual trade ω_3 in a matching maximum score inequality (10) based on swapping windshield suppliers.

The car parts data come from SupplierBusiness, an analyst firm. I merge them with car sales data collected from several sources for the United States and several large countries in Western Europe. I focus on 30 large component categories, such as air conditioning parts, body parts and transmission parts. In the merged and cleaned data, there are 941 suppliers, 11 assemblers (parent companies), 46 car brands, 260 car models, and 34,836 car parts. While the data cover different model years, for simplicity I ignore the time dimension and treat each market as clearing simultaneously.[19] I treat each component category as a statistically independent matching market.[20] Therefore, I use the matching maximum score objective function for multiple markets in (15).[21]

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.[22] 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 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.[23] 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 maximum score estimator. The matching markets modeled here contain many more agents than the markets modeled in many non-marriage 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. I focus on specialization in the portfolio of matches for suppliers and

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

[20]The same supplier may appear in multiple component categories, and so a researcher might want to model spillovers across component categories. Pooling component categories into one large market creates no new issues with the AH model or the matching maximum score estimator. 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.

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

[22]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.

[23]A few suppliers are owned by assemblers. I ignore the 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).

assemblers. Along with my related use of the estimator introduced here in Fox and Bajari (2013), an earlier draft of the current paper was the first empirical application to a many-to-many matching market where the valuation from a set of matches (or trades) is not additively separable across the individual matches. Finally, the prices of the car parts are not in publicly available data. The matching estimator does not require data on the prices of trades, even though prices are present in the economic model being estimated.

6 Costs of Assemblers Divesting Brands

6.1 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.[24] Opel is based in Germany and Vauxhall is based in the United Kingdom. Consistent with the close link between Opel and Vauxhall, they will be grouped together as one brand, Opel, in the empirical work. Over the period of the data, Opel also had assembly plants in Belgium, Hungary, Poland, and Russia.

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 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 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.

6.2 Valuation Functions of Observable Types

This section estimates the parameters in the valuation functions over observable types for assemblers and suppliers for the portfolio of car part trades each firm buys or sells. Let the notation for the observable type j^s emphasize that the firm in question is a supper or seller and the notation j^b emphasize that the firm is a buyer or assembler, as the car parts industry is an explicitly two-sided market. I use the functional forms $\pi_\theta\left(j^b, \Phi\right) = X\left(j^b, \Phi\right)' \theta^b$ for buyers and $\pi_\theta\left(j^s, \Psi\right) = X\left(j^s, \Psi\right)' \theta^s$

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

for sellers, with $\theta = \left(\theta^b, \theta^s\right)$. The elements of the vectors $X\left(j^b, \Phi\right)$ and $X\left(j^s, \Psi\right)$ are measures of how specialized each portfolio of car part trades is at several levels.

6.2.1 Valuation Functions for Suppliers

For suppliers, $X\left(j^s, \Psi\right)$ tracks specialization 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 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 the corresponding parent group scalar element $X_{\mathrm{PG}}\left(j^s, \Psi\right)$ of the vector $X\left(j^s, \Psi\right)$ is

$$X_{\mathrm{PG}}\left(j^s, \Psi\right) = \left(\frac{\#\,\text{Chrysler parts in}\,\Psi}{\#\,\text{total parts in}\,\Psi}\right)^2 + \left(\frac{\#\,\text{Ford parts in}\,\Psi}{\#\,\text{total parts in}\,\Psi}\right)^2 + \left(\frac{\#\,\text{GM parts in}\,\Psi}{\#\,\text{total parts in}\,\Psi}\right)^2. \tag{16}$$

As this specialization measure enters the valuation function for a supplier, $X_{\mathrm{PG}}\left(j^s, \Psi\right)$ 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 differs from antitrust; here the HHI is a measure of specialization for a portfolio Ψ of car part trades for a particular supplier j^s and is not a measure of concentration in the overall industry for car parts. The specialization measure $X_{\mathrm{PG}}\left(j^s, \Psi\right)$ can be computed both for the trades Ψ for a supplier in the data and in the counterfactual trades in a matching maximum score inequality (10).[25]

When I consider the counterfactual of GM divesting Opel and making it an independent assembler or parent company, the changes in total valuation 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.

The pattern of sorting across trades 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, 2012).

By construction, two parts for the same car model 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 $X\left(j^s, \Psi\right)$ are highly correlated. Just as univariate linear least

[25] Many other upstream firm characteristics would be endogenous at the level of the competitive equilibrium 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. Plant location could be added as an extra element to a trade ω in other work.

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 measure in $X(j^s, \Psi)$ separately will be inadequate here. A univariate analysis of say $X_{\text{PG}}(j^s, \Psi)$ would just amount to saying that the corresponding element of θ is positive when each supplier does more business with certain parent groups than others. In principle, even this conclusion about the sign of the parameter 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?

6.2.2 Valuation Functions for Assemblers

The valuation function of assemblers has a similar functional form, focusing on specializing in a small number of suppliers. Let Φ be a portfolio of car part trades for buyer or assembler j^b. I consider parent group, brand and model specialization in the vector $X(j^b, \Phi)$. For conciseness, I do not include a term for specialization at the continent of brand headquarters level.

Consider the Herfindahl index for the concentration of suppliers selling parts to an assembler. Given a portfolio Φ, let $s(\Phi)$ be the set of distinct suppliers who sell at least one car part trade in Φ. Then define the scalar

$$X_{\text{PG}}\left(j^b, \Phi\right) = \sum_{i \in s(\Phi)} \left(\frac{\#\,\text{trades sold by supplier } i \text{ in } \Phi}{\#\,\text{total trades in } \Phi} \right)^2 .$$

Next, $X_{\text{Br}}\left(j^b, \Phi\right)$ is the mean of such a Herfindahl index computed for each brand separately. Consider GM and say that the only two brands of GM are Chevrolet (Chevy) and Opel and let $s(\Phi, \text{Opel})$ be the set of suppliers selling parts to Opel in Φ. Then, for GM,

$$X_{\text{Brand}}\left(j^b, \Phi\right) = \frac{1}{2} \sum_{i \in s(\Phi, \text{Opel})} \left(\frac{\#\,\text{trades sold by supplier } i \text{ to Opel in } \Phi}{\#\,\text{total trades for Opel in } \Phi} \right)^2 + \frac{1}{2} \sum_{i \in s(\Phi, \text{Chevy})} \left(\frac{\#\,\text{trades sold by supplier } i \text{ to Chevy in } \Phi}{\#\,\text{total trades for Chevy in } \Phi} \right)^2 . \quad (17)$$

Likewise, $X_{\text{Model}}\left(j^b, \Phi\right)$ is the mean across car models sold by GM of the Herfindahl index calculated for the sellers of parts to each car model separately. As with suppliers, the elements of $X\left(j^b, \Phi\right)$ can be computed for the counterfactual trades in the matching maximum score inequalities (10).

The matching maximum score inequalities used in estimation keep the number of car part trades sold 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 matching maximum score inequalities where each car part and each car model are weighted equally. This paper models the car parts industry, 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 matching maximum score inequalities used in estimation do not identify a pure scale economy for GM owning Opel. Instead, the matching maximum score inequalities focus on the gains to assemblers and particularly to suppliers from specialization, for a fixed number of car part trades.

6.3 Estimates for Valuation Functions

Table 1 presents the point estimates and confidence intervals for the parameter vector θ in the valuation functions for observable types, for both assemblers and suppliers. I randomly sample a maximum of 10,000 matching maximum score inequalities (10) per component category. All theoretically valid inequalities with two different suppliers are sampled with an equal probability. I use the set identified subsampling procedure of Romano and Shaikh (2008) to construct confidence regions. See Appendix B for details on estimation and inference.

The parameter on assembler (parent group) specialization for suppliers is normalized to be ± 1. The other parameters in Table 1 are interpreted relative to the parameter on parent group specialization. One finding in that the point estimates of the assembler parameters have a much lower order of magnitude than the supplier parameters and the parameters have wide confidence bands, always including 0. For assemblers, the upper bounds of the confidence bands for parent group, brand and model specialization are lower than the lower bounds for the analogous specialization measures for suppliers. Therefore, for these specialization measures one can at least statistically conclude that supplier specialization measures are more important. This difference between the point estimates for assemblers and suppliers is not because of a difference in the units of $X\left(j^b, \Phi\right)$ and $X\left(j^s, \Psi\right)$; the rightmost columns of Table 1 report the means and standard deviations of the specialization measures for realized matches for both suppliers and assemblers. The specialization (HHI) measures are about the same magnitudes for both suppliers and assemblers. 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 the valuation functions. One caveat is that the confidence intervals for assembler specialization at the model levels do contain larger, in absolute value, coefficient magnitudes.

For suppliers, Table 1 shows that all four coefficients on supplier specialization are positive, meaning as expected specialization on these dimensions increases the valuation of suppliers. The point estimates show that a given level of specialization at the parent-group level is about as important in valuation as the same level of specialization at the continent-of-brand-headquarters level. At the same time, the standard deviation of parent-group-specialization HHI, across realized matches, is 0.18, meaning the variation in parent-group specialization across suppliers is lower than for some other specialization measures. A naive researcher might be inclined to interpret this level of dispersion as evidence parent-group specialization is unimportant. This would be wrong: the matching 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 valuation from a set of trades.

Table 1 also shows that supplier specialization at the brand and model levels is even more important than specialization at the parent-group level, as the brand and model confidence intervals do not contain +1. The high point estimate of 376 for model specialization possibly 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 (0.26–0.28) for the continent, brand and model measures, with parent group specialization being a little lower at 0.18. Again, naive researchers might use the HHI means to conclude that specialization at the model level is less important or use the standard deviations to conclude that specialization at the continent, brand and model levels are equally important. The estimates of the valuation functions give statistically consistent estimates of the relative importance of the types of specialization in the valuation functions for supplier relationships.

Table 1 also shows that there are 298,272 inequalities used in estimation. Of those, 82% are satisfied at the reported point estimates. The fraction of satisfied inequalities is a measure of statistical fit. Appendix C presents estimates where the HHI specialization measures use different weighting schemes, including weighting schemes using data on car model sales in Europe and North America. The specifications in Appendix C result in lower numbers of inequalities being satisfied at the parameter estimates. Therefore, these alternatives result in statistically worse fit and so are not presented in the main text. However, a common finding in Appendix C is that the assembler parameters become more important, particularly the parameter on assembler model specialization.

6.4 Supplier Valuation 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 industry is not necessarily informative about assembler economies of scale and scope.

Using information from the car parts industry, 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 prices of trades: suppliers will charge higher prices to GM for car parts. In each component category, I construct the counterfactual sum of valuations from observable types to suppliers if Opel and the rest of GM are now treated as separate assemblers, or parent groups. The same sellers supply the same car part trades to the same car models, but now the Opel models are produced by an independent parent group. In (16), some car 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 the parent group specialization measure $X_{\mathrm{PG}}(j^s, \Psi)$ times its estimated parameter θ^s_{PG} gives the decrease in the valuation for each supplier who sells at least one part to Opel. I focus on a percentage decrease measure

$$\frac{\theta^s_{\mathrm{PG}} \Delta X_{\mathrm{PG}}\left(j^s, \Psi\right)}{X\left(j^s, \Psi\right)' \theta},$$

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

Table 2 reports statistics for the distribution of percentage changes in valuation 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 tiny, at 0.04%. The other quantiles are tiny as well. This partly 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. This result also follows from the parameter estimates in Table 1, where the point estimates for the coefficients on brand and especially model specialization are many times larger than the coefficient on parent group specialization.[26]

6.5 International Trade Application

Appendix D contains estimates of a separate specification motivated by how barriers to international trade may upgrade the quality of domestic suppliers.

7 Conclusions

This paper discusses the estimation of valuation functions in matching games with transferable utility. A matching maximum score estimator is introduced for the matching with trades model of Azevedo and Hatfield (2015, Section 6), which has many special cases of empirical relevance. The matching maximum score estimator is computationally simple and semiparametric.

The empirical work answers two policy questions surrounding the automotive industry. First, the paper estimates the relative loss in valuation 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 model level, which is not affected by the divestment, are higher than the gains to specialization at the parent group level. Second and in an appendix, the paper estimates the gain to, say, North American suppliers from the presence of Asian-based assemblers in North America. Both estimates are inferred from a new type of data, the portfolios of car part trades from each supplier.

[26] I compute but do not report the small changes in GM's and Opel's valuations from divesting Opel. Because the coefficient estimates on assembler specialization in Table 1 are small in magnitude, the overwhelming effect in profit levels is estimated to be on suppliers.

A Proofs

A.1 Proposition 1: Rank Order Property

The proof cites Graham (2011, Theorem 4.1). His theorem is stated for one-to-one, two-sided matching or marriage and, in his notation, considers two observable types of men k and m and two observable types of women l and n. Graham's proof works by considering the so-called suballocation of the two observable types of men and the two observable types of women, the suballocation $klmn$ in his notation. In the suballocation $klmn$, Graham considers the probability within the suballocation of k and l matching, which is called r^{klmn}. The probability of k matching with n is $p^{klmn} - r^{klmn}$, the probability of m matching with l is $q^{klmn} - r^{klmn}$, and the probability of m matching with n is $1 - p^{klmn} - q^{klmn} + r^{klmn}$.

The statement that drawing two matches between k and l and m and n are more likely than drawing two matches between k and n and m and l is

$$\left(1 - p^{klmn} - q^{klmn} + r^{klmn}\right) r^{klmn} > \left(p^{klmn} - r^{klmn}\right)\left(q^{klmn} - r^{klmn}\right). \tag{18}$$

Algebra at the end of Section 4.3.2 in Graham, after correcting typos, shows that this inequality is equivalent to $r^{klmn} > p^{klmn} q^{klmn}$. The conclusion of Theorem 4.1 on the same page of Graham can be rewritten to state that

$$\delta_{mn} + \delta_{kl} \geq \delta_{ml} + \delta_{kn} \tag{19}$$

if and only if $r^{klmn} > p^{klmn} q^{klmn}$, which as just said is equivalent to the inequality (18). The inequality (19) can be seen as a matching maximum score inequality (10). Therefore, Graham (2011, Theorem 4.1) is a rank order property for one-to-one, two-sided matching without data on prices.

The setup in the AH model is more general than the model of one-to-one, two-sided matching or marriage in Graham. However, the extra generality can be handled by conditioning. In an matching maximum score inequality $g \in G$ in my notation, there are the two observable agent types of buyers and the two observable agent types of sellers in $H_g = \{b(\omega_1), s(\omega_1), b(\omega_2), s(\omega_2)\}$. While the AH model does not necessarily assign roles of buyers and sellers ex ante, a matching maximum score inequality (10) does condition on these roles on the left side of the inequality and the right side of the inequality. A complication is the observable type $b(\omega_1)$ that is the buyer on ω_1 on the left side of the inequality could be a seller on, say, trade ω_3 on the right side of the inequality. This switch of roles of buyer and seller does not change the proof in Graham.

Likewise, agents in the AH model make sets of trades (Φ, Ψ). This can be handled by conditioning on the set of trades (Φ_j, Ψ_j) other than ω_1–ω_4 for all four observable agent types in $j \in H_g$. In other words, condition on the joint event C that $b(\omega_1)$ picks either $\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}\right)$ or $\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right)$, $s(\omega_1)$ picks either $\left(\Phi_{s(\omega_1)}, \Psi_{s(\omega_1)}\right)$ or $\left(\bar{\Phi}_{s(\omega_1)}, \bar{\Psi}_{s(\omega_1)}\right)$, and similarly for $b(\omega_2)$ and $s(\omega_2)$.

In what follows, abbreviate $\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right)$ with $\bar{A}^j\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g\right)$. When forming choice probabilities conditional on the joint event C mentioned just above, the conditional choice probabilities for say $b(\omega_1)$ will multiplicatively factor into $\bar{A}^{b(\omega_1)}(\Phi, \Psi \mid g)$ and three choice probabilities for the other three agents. In an inequality such as the probability statement (14) in the statement of the proposition, the choice probabilities for the other three agents are the same multiplicative correction for $\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g\right)$ and $\bar{A}^{b(\omega_1)}\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)} \mid g\right)$ and cancel out on either side of the inequality. The choice probabilities multiplicatively factor as the choice probabilities of agents are mutually independent conditional on the observable type $j \in J$ in the continuum AH model.

Divide, on the left and right sides of (18), $\left(1 - p^{klmn} - q^{klmn} + r^{klmn}\right)$ and $\left(q^{klmn} - r^{klmn}\right)$ by the constant

$$\left(1 - p^{klmn} - q^{klmn} + r^{klmn}\right) + \left(q^{klmn} - r^{klmn}\right).$$

This changes those terms into, in my notation, the conditional choice probabilities $\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g\right)$ and $\bar{A}^{b(\omega_1)}\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)} \mid g\right)$, respectively, after canceling out the multiplicatively factorable choice probabilities for the three agents other than $b(\omega_1)$ for conditioning on the joint event C, as just discussed. A similar argument applies to give $\bar{A}^{b(\omega_2)}\left(\Phi_{b(\omega_2)}, \Psi_{b(\omega_2)} \mid g\right)$ and $\bar{A}^{b(\omega_2)}\left(\bar{\Phi}_{b(\omega_2)}, \bar{\Psi}_{b(\omega_2)} \mid g\right)$. Therefore, the probability statement (18) becomes the probability statement (14) in the statement of the proposition. Hence, we have proved the rank order property for the AH model without data on the prices of trades.

A.2 Theorem 1: Set Identification

In what follows, abbreviate $\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid \Phi_{b(\omega_1)}, \Psi_{b(\omega_1)}, \bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)}\right)$ with $\bar{A}^j\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g\right)$. By the law of iterated expectations, the expectation of the matching maximum score objective function (11) can be written

$$\frac{1}{2}\sum_{g_1 \in G} \Pr\left[g_1 \text{ or } g_2(g_1) \text{ included}\right] \cdot \Big\{ \bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g_1\right) \cdot \bar{A}^{b(\omega_2)}\left(\Phi_{b(\omega_2)}, \Psi_{b(\omega_2)} \mid g_1\right) \cdot 1\left[Z_{g_1}'\theta \geq 0\right] +$$
$$\bar{A}^{b(\omega_1)}\left(\bar{\Phi}_{b(\omega_1)}, \bar{\Psi}_{b(\omega_1)} \mid g_1\right) \cdot \bar{A}^{b(\omega_2)}\left(\bar{\Phi}_{b(\omega_2)}, \bar{\Psi}_{b(\omega_2)} \mid g_1\right) \cdot 1\left[Z_{g_1}'\theta < 0\right] \Big\}.$$

By a requirement in the main text, if $g_1 \in G$, then $g_2(g_1) \in G$. The 1/2 is to remove double counting: counting g_1 once as g_1 in the summation as once as $g_2(g_1')$ for $g_1' = g_2(g_1)$. The calculation $\Pr\left[g_1 \text{ or } g_2(g_1) \text{ included}\right]$ is over the four observable types of i_1–i_4 in (11) and all the trades except ω_1–ω_4 in inequality g_1 and, hence, the reverse direction inequality $g_2(g_1)$. Further, one of those actually enters with a strict $>$ instead of the weak $\geq$. So the forward direction inequality $Z_{g_1}'\theta \geq 0$ is mutually exclusive with $Z_{g_1}'\theta < 0$. Only one of the two inequalities can enter the objective function with nonzero weight for a given parameter θ. By the rank order property in Proposition 1, the maximum of the two products of conditional choice probabilities of the form

$$\bar{A}^{b(\omega_1)}\left(\Phi_{b(\omega_1)}, \Psi_{b(\omega_1)} \mid g_1\right) \cdot \bar{A}^{b(\omega_2)}\left(\Phi_{b(\omega_2)}, \Psi_{b(\omega_2)} \mid g_1\right)$$

will be included at the true parameter value for θ. Any other parameter value θ results in either the same objective function value or a lower objective function value where some smaller weight contributes to the objective function value. Therefore, the matching maximum score objective is (perhaps not uniquely) globally maximized at the true parameter value.

B Estimation and Inference for the Car Parts Empirical Work

I first describe point estimation and then inference. I sample 10,000 inequalities for each of 29 matching markets given by component categories. For the 30th component category, I use the maximum number of inequalities, 8,272, for that market. To create one of the 10,000 inequalities for a market, I randomly sample two car parts from different suppliers and include the resulting matching maximum score inequality where the two car parts are exchanged between the suppliers.

I use the numerical optimization routine differential evolution, in Mathematica. For differential evolution, I use a population of 250 points and a maximum of 15,000 iterations. The numerical optimization is run 20 times with different initial populations of 250 points. For the specification in Table 1, all 20 runs found the same objective function value and the same point estimates, up to two significant digits.

In the maximum score objective function, an inequality is satisfied if the left side plus the constant $+0.0000000001$ exceeds the right side. This small perturbation to the sum of profits on the left side ensures that inequalities such as $0 > 0$ are counted as being satisfied consistently, rather than inconsistently because of some numerical-approximation error resulting in, say, $2.0 \times 10^{-15} > 1.0 \times 10^{-15}$.

I use the inference procedure of Romano and Shaikh (2008), which is valid under both set identification and point identification. Let Θ_0 be the identified set for the parameter θ. The authors provide a subsampling procedure to construct a confidence region $\mathscr{C}$ called, by those authors, a "confidence region for the identifiable parameters that is uniformly consistent in level", under the conditions of their Theorem 3.3. This definition of the properties of a confidence region is equation (3) of their paper. Under easier to verify conditions, their procedure produces "confidence regions for identifiable parameters that are pointwise consistent in level", or their equation (2). Computationally, I implement equations (12) and (10) in their paper.

I have 30 component categories. I treat these component categories as separate markets for the industry-specific reasons discussed in the main text. I perform asymptotics in the number of suppliers in the market, as this is an explicitly two-sided market where a supplier cannot also be an assembler, as in Example 2. Subsampling requires a choice of subsample size. Unfortunately, the literature has not provided a data driven method to pick this tuning parameter. I use 33% of the suppliers from each of the 30 component categories. Given each subsampled set of suppliers, that subsample uses only matching maximum score inequalities where both suppliers whose valuation functions are in the inequality are in the subsample. I use 500 subsamples; in experiments results are robust when using more subsamples.

In the car parts empirical work, the maximum score or maximum rank correlation objective function is, for $D = 30$ component categories,

$$Q(\theta) = \binom{N}{2}^{-1} \sum_{d=1}^{D} \sum_{i_1=1}^{N_d - 1} \sum_{i_2 = i_1 + 1}^{N_d} \sum_{g \in G^d_{i_1, i_2}} 1\left[Z'_{g,d} \theta \geq 0 \right],$$

where $G^d_{i_1,i_2}$ is the set of included matching maximum score inequalities where i_1 and i_2 exchange one car part each. Romano and Shaikh (2008) is written where one minimizes a function and where the population objective function's minimum value is 0. In maximum score, one maximizes the objective function value and the value of the population objective function can only be estimated using the finite sample objective function. Following the suggestion at the beginning of their Section 3.2, I work with the objective function $-Q(\theta) - (-\max_\theta Q(\theta))$.

I define the "sample size" for the entire procedure to be $N = \max_{d \in D} N_d$. This choice is arbitrary and does not impact the reported confidence regions. I implement the procedure of Romano and Shaikh (2008) using the rate of convergence for the objective function of $\sqrt{N}$, following the results of Sherman (1993) for the case of point identified maximum rank correlation estimators.

The confidence regions reported in tables are projections onto the axes of the confidence region $\mathscr{C}$ from Romano and Shaikh (2008). Computationally, the lower bound for the confidence region for a scalar parameter θ_1 is found by minimizing the parameter θ_1 subject to the constraint that the entire parameter vector θ is in $\mathscr{C}$. Likewise, the upper bound for the confidence region θ_1 is found by maximizing the parameter θ_1 subject to the entire parameter vector being in $\mathscr{C}$.

C Alternative Weighting Schemes for Specialization Measures

This appendix discusses versions of the estimates in Table 1 where the included HHI specialization measures use different weighting schemes, many involving data on car sales for car models in Europe and the United States. These alternative weighting schemes result in worse statistical fit than the estimates in Table 1, in the sense that fewer inequalities are satisfied. For this reason, I report the estimates in this appendix instead of the main text.

In terms of sales, car models primarily sold in Europe are matched to European sales from Western Europe and car models primarily from North America are matched to sales from the United States. Note that all weighting schemes affect the HHI specialization measures. I do not explore weighting the maximum score inequalities in the objective function, although that could be pursued.

Let w_c be a weight for a particular car part c. The example parent group supplier HHI measure in (16) becomes

$$X_{\text{PG}}\left(j^s, \Psi\right) = \left(\frac{\text{sum Chrysler weights in }\Psi}{\text{sum all weights in }\Psi}\right)^2 + \left(\frac{\text{sum Ford weights in }\Psi}{\text{sum all weights in }\Psi}\right)^2 + \left(\frac{\text{sum GM weights in }\Psi}{\text{sum all parts in }\Psi}\right)^2.$$

Analogous schemes are used for other supplier HHI specialization measures and for assembler specialization at the parent group level. The example brand assembler HHI measure in (17) becomes

$$X_{\text{Brand}}\left(j^b, \Phi\right) = \bar{w}_{\text{Opel}} \sum_{i \in s(\Phi, \text{Opel})} \left(\frac{\text{sum weights sold by supplier } i \text{ to Opel in } \Phi}{\text{sum weights for Opel in } \Phi}\right)^2 + \bar{w}_{\text{Chevy}} \sum_{i \in s(\Phi, \text{Chevy})} \left(\frac{\text{sum weights sold by supplier } i \text{ to Chevy in } \Phi}{\text{sum weights for Chevy in } \Phi}\right)^2,$$

where, for example,

$$\bar{w}_{\text{Chevy}} = \frac{\text{sum weights for Chevy in GM}}{\text{sum weights in GM}}.$$

Note that the brand weights like $\bar{w}_{\text{Chevy}}$ are not recomputed for counterfactual trades Φ, as indeed Chevrolet and other brands still produce the same car models with the same need for car parts in the counterfactuals in matching maximum score inequalities.

Table 3 reports the estimates with weights. There are three specifications, each with different weights. Note that the parameters multiply different explanatory variables across the three specifications, so the parameters have slightly different interpretations. However, the means and standard deviations of the weighted HHI measures are qualitatively similar to those reported in Table 1 for the baseline specification.

The first weighting scheme in Table 3 does not use sales data; it sets

$$w_c = \frac{1}{\#\,\text{car parts for model in market } d}.$$

This scheme weights a car part on a car model with more parts in a component category less, in order to equalized the contribution of car models to HHI calculations. The second weighting scheme uses the weights

$$w_c = \text{sales of car model}.$$

The third weighting scheme combines the previous two, as in

$$w_c = \frac{\text{sales of car model}}{\#\,\text{car parts for model in market } d}.$$

Compared to Table 1, Table 3 shows less of a role for supplier model specialization and much more of a role for assembler model HHI specialization. In all three of the specifications, the lower bound for the confidence region for supplier HHI model specification is negative and large in magnitude relative to the +1 normalization for supplier parent group specialization. The point estimate is negative in two of the three specifications. In all three specifications, the upper bound of the confidence region for assembler model specialization is unbounded. The point estimate for assembler model specialization is unbounded in two specifications and still quite large, at 202, for the other specification.

Also compared to Table 1, there is now statistical uncertainty about the sign of supplier brand specialization, although the point estimate is about the same as for parent group specialization, which is normalized to +1. Similarly to Table 1, 0 is still in the confidence regions for assembler parent group and brand specialization.

Overall, I emphasize Table 1 in the main text because of its higher statistical fit.

D 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 also assemble cars in Europe and North America. While some car 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 higher-quality cars must typically also be of higher 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 might be consistent with suppliers to Asian assemblers being higher quality than 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 supplier sells 30 of them, its quality measure will be 0.30. In notation, one aspect of an observable firm type j^s for a supplier is

$$j^s_{\text{Asia}} = \frac{\#\,\text{Asian assembler parts supplied}}{\text{total}\,\#\,\text{Asian assembler parts all suppliers}}.$$

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 j^s_{Asia} is treated as an aspect of observable firm type j^s of a supplier. If it were recomputed for new portfolios Φ without interactions in a valuation function, it would difference out of the matching maximum score inequalities (10). Instead, the vector $X(j^s, \Psi)$ contains a new element that is the interaction of the above Asian quality measure with specialization by the continent headquarters of the brand, discussed earlier:

$$X_{\text{AsiaCont}}(j^s, \Psi) = j^s_{\text{Asia}} \cdot X_{\text{Cont}}(j^s, \Psi).$$

The interpretation of the corresponding supplier parameter in θ, if it is estimated to be negative, is that suppliers with higher j^s_{Asia} (greater shares of the industry for supplying Asian assemblers) gain less benefit from selling car parts to only one continent of assembler than suppliers with lower j^s_{Asia}. Thus, suppliers 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 might be that suppliers with high j^s_{Asia} have diverse (across continents of assembler origin) portfolios of car parts that they supply. This diversity might be interpreted as a sign of quality.

Even if the parameter on $X_{\text{AsiaCont}}(j^s, \Psi)$ 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 j^s_{Asia} 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 j^s_{Asia} as an economically exogenous characteristic, rather than recomputing the Asian market share for counterfactual sets of trades Φ in the right sides of matching maximum score inequalities. I have explored the specification where notationally j^s_{Asia} is replaced by $X_{\text{Asia}}(j^s, \Psi)$, which is recomputed for counterfactual sets of trades Ψ. The corresponding point estimate on the interaction is ≈ 0, with a wide confidence interval in terms of economic magnitudes. An explanation for the point estimate close to 0 is that a new effect is introduced to the model: the inequalities ask why more suppliers 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 j^s_{Asia}.

Table 4 presents the point estimates from the preferred specification. The other covariates are the assembler and supplier specialization measures in Table 1, which have similar point estimates. The scale normalization is still on parent group specialization. With the interaction term $X_{\text{AsiaCont}}(j^s, \Psi)$ involving continent specialization, the normalization can only be understood by substituting a typical value for j^s_{Asia} into the interaction term $X_{\text{AsiaCont}}(j^s, \Psi)$ and comparing also the coefficient on continent specialization without an interaction.

The new addition to Table 4 is the estimate on the interaction term $X_{\text{Asia}}(j^s, \Psi)$, which would use an estimated 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 possibly can win business from non-Asian assemblers. The estimate of the parameter on $X_{\text{Asia}}(j^s, \Psi)$ is -0.261 and the mean and standard deviation of j^s_{Asia}, not listed in the table, are 0.069 and 0.102, respectively. Therefore, a one-standard deviation change in j^s_{Asia} creates a change of $-0.261 \cdot 0.102 = -0.0266$ 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 0.171, will have a total coefficient on continent-of-brand specialization of $+1.03 - 0.261 \cdot 0.171 = 0.985$, or approximately 1. This is a small magnitude change. Now that the confidence region for the interaction parameter is huge, from -30 to

32. The data do not pin down this effect.

With a bigger in absolute value and more precisely estimated effect, the interpretation would have been 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 assemblers from other continents. Thus, the evidence from sorting in the market for car parts would have suggested 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 would have suggested that having higher quality assemblers in Europe and North America raises the quality of suppliers. Thus, in the automotive industry there might have been evidence of indirect benefits to domestic suppliers and assemblers from the trade barriers that encourage Asian assemblers to locate in Europe and North America.

References

Agarwal, Nikhil and William Diamond, "Identification and Estimation in Two-Sided Matching Markets," 2013. working paper.

Akkus, Oktay, J. Anthony Cookson, and Ali Hortacsu, "The Determinants of Bank Mergers: A Revealed Preference Analysis," *Management Science*, forthcoming.

Azevedo, Eduardo M. and John William Hatfield, "Existence of Equilibrium in Large Matching Markets with Complementarities," 2015. University of Pennsylvania working paper.

Baccara, Mariagiovanna, Ayse Imrohoroglu, Alistair Wilson, and Leeat Yariv, "A Field Study on Matching with Network Externalities," *American Economic Review*, 2012, *102* (5).

Bajari, Patrick, 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," *Journal of Labor Economics*, 2013, *31* (1), 83–117.

Briesch, Richard A., Pradeep C. Chintagunta, and Rosa L. Matzkin, "Semiparametric estimation of brand choice behavior," *Journal of the American Statistical Association*, December 2002, *97* (460), 973–982.

Chen, Xiaohong, "Large Sample Sieve Estimation of Semi-Nonparametric Models," in "Handbook of Econometrics," Elsevier Science, 2007.

___ , **Elie Tamer, and Alexander Torgovitsky**, "Sensitivity Analysis in Semiparametric Likelihood Models," 2011. Working paper.

Chiappori, Pierre-André and Bernard Salanié, "The Econometrics of Matching Models," *Journal of Economic Literature*, forthcoming.

___ , ___ , **and Yoram Weiss**, "Partner Choice and the Marital College Premium: Analyzing Marital Patterns Over Several Decades," 2015. Columbia University working paper.

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

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

Delgado, M.A., J.M. Rodriguez-Poo, and M. Wolf, "Subsampling inference in cube root asymptotics with an application to Manski's maximum score estimator," *Economics Letters*, 2001, *73* (2), 241–250.

Dupuy, Arnaud and Alfred Galichon, "Personality traits and the marriage market," *Journal of Political Economy*, 2014, *122* (6), 1271–1319.

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.

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, *1* (2), 203–254.

___ **and Patrick Bajari**, "Measuring the Efficiency of an FCC Spectrum Auction," *American Economic Journal: Microeconomics*, 2013, *5* (1), 100–146.

___ , **David Hsu, and Chenyu Yang**, "Unobserved Heterogeneity in Matching Games with an Application to Venture Capital," 2015. Rice University working paper.

Gale, David, *The Theory of Linear Economic Models*, McGraw-Hill, New York, 1960.

Galichon, Alfred and Bernard Salanie, "Cupid's Invisible Hand: Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," 2012. Ecole Polytechnique working paper.

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

Graham, Bryan S., "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in J. Benhabib, A. Bisin, and M. Jackson, eds., *Handbook of Social Economics*, Elsevier, 2011, chapter 19.

_ , "Errata in "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers"," 2013. Online erratum.

Hajivassiliou, Vassilis A and Daniel L McFadden, "The method of simulated scores for the estimation of LDV models," *Econometrica*, 1998, pp. 863–896.

Han, A.K., "Non-parametric analysis of a generalized regression model," *Journal of Econometrics*, 1987, *35*, 303–316.

Hatfield, John William, Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp, "Stability and Competitive Equilibrium in Trading Networks," *Journal of Political Economy*, 2013, *121* (5), 966–1005.

Heckman, James J., Rosa L. Matzkin, and Lars Nesheim, "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica*, September 2010, *78* (5), 1569–1591.

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

Kim, J. and D. Pollard, "Cube root asymptotics," *The Annals of Statistics*, 1990, pp. 191–219.

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

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

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.

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

Manski, C.F., "Semiparametric analysis of discrete response:: Asymptotic properties of the maximum score estimator," *Journal of Econometrics*, 1985, *27* (3), 313–333.

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

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

McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, 1973, pp. 105–142.

McFadden, Daniel L., "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometricac*, September 1989, *57* (5), 995–1026.

Menzel, Konrad, "Large matching markets as two-sided demand systems," *Econometrica*, 2015, *83* (3), 987–941.

Mindruta, Denisa, Mahka Moeen, and Rajshree Agarwal, "A Two-Sided Matching Approach for Partner Selection and Assessing Complementarities in Partners' Attributes in Inter-Firm Alliances," *Strategic Management Journal*, 2015.

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, W.K. and D. McFadden, "Large Sample Estimation and Hypothesis Testing," in "Handbook of Econometrics," Vol. 4, Elsevier, 1994, pp. 2111–2245.

Novak, Sharon and Birger Wernerfelt, "On the Grouping of Tasks into Firms: Make-or-Buy with Interdependent Parts," *Journal of Economics and Management Strategy*, 2012, *21* (1), 53–77.

_ **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.

Pakes, A. and D. Pollard, "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 1989, *57* (5), 1027–1057.

Romano, Joseph P and Azeem M Shaikh, "Inference for identifiable parameters in partially identified econometric models," *Journal of Statistical Planning and Inference*, 2008, *138* (9), 2786–2807.

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, *82* (1), 34–55.

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.

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

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.

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

Uetake, Kosuke and Yasutora Watanabe, "Entry by Merger: Estimates from a Two-Sided Matching Model with Externalities," 2012. Northwestern University working paper.

Table 1: Specialization By Suppliers and Assemblers

	Valuation Function Estimates		Sample Statistics for HHI Measures	
HHI Measure	Point Estimate	95% CI Set Identified	Mean	Standard Deviation
Suppliers				
Parent Group	+1	Superconsistent	0.35	0.28
Continent	1.04	(0.0482,9.45)	0.76	0.18
Brand	23.9	(1.29,121)	0.25	0.27
Model	376	(278,933)	0.17	0.26
Assemblers				
Parent Group	-0.007	(-1.30,0.202)	0.14	0.11
Brand	-0.005	(-1.99,0.705)	0.35	0.33
Model	-0.003	(-3.36,33.5)	0.58	0.60
# Inequalities	298,272			
% Satisfied	82.3%			

The parameter on parent group specialization is fixed at +1. Estimating it with a smaller number of inequalities always finds the point estimate of +1, instead of −1. The estimate of a parameter that can take only two values is superconsistent, so I do not report a confidence interval. See Appendix B for details on estimation and inference.

Table 2: Percentage Valuation Change By Suppliers From GM Divesting Opel

Quantile	
0	-0.0032
0.10	-0.0014
0.25	-0.0008
0.50 (median)	-0.0004
0.75	-0.0002
0.90	-0.00008
1	~0

This table uses the point estimates from Table 1 to calculate the valuations from observable types of suppliers before and after GM divests Opel. In the model, Opel becomes a separate parent group. For each supplier, selling one or more parts to Opel and one or more cars to another GM brand, I calculate $\frac{\theta^s_{\mathrm{PG}} \Delta X_{\mathrm{PG}}(j^s, \Psi)}{X(j^s, \Psi)'\theta}$. Each supplier that operates in multiple component categories (markets) is treated separately in each component category.

Table 3: Specialization By Suppliers and Assemblers with Different Weighting Schemes

Weights→	1 Divided by #Parts		Sales		Sales Divided by #Parts	
HHI Measure	Point Estimate	95% CI	Point Estimate	95% CI	Point Estimate	95% CI
Suppliers						
Parent Group	+1	Supercon.	+1	Supercon.	+1	Supercon.
Continent	0.982	(0.405,4.66)	0.733	(0.582,1.29)	0.717	(0.298,3.12)
Brand	0.988	(-0.154,8.52)	0.967	(0.158,3.19)	0.894	(-0.52,6.06)
Model	0.0806	(-8.79,2.52)	-1.97	(-4.14,-1.31)	-1.88	(-6.18,-0.374)
Assemblers						
Parent Group	0.474	(-0.0902,4.70)	≈ 0	(-0.0318,0.0654)	-0.0448	(-0.36,0.0333)
Brand	-0.380	(-3.77,1.79)	≈ 0	(-0.0435,0.556)	0.0462	(-0.212,1.2)
Model	202	(100,$+\infty$)	$+\infty$	($+\infty$, $+\infty$)	$+\infty$	($+\infty$, $+\infty$)
# Inequalities	298,272		298,272		298,272	
% Satisfied	76.3%		74.4%		74.3%	

The parameter on parent group specialization is fixed at +1. Estimating it with a smaller number of inequalities always finds the point estimate of +1, instead of −1. The estimate of a parameter that can take only two values is superconsistent, so I do not report a confidence interval. See Appendix B for details on estimation and inference.

Table 4: Supplier Competitive Advantages From Asian Assemblers

	Valuation Function Estimates	
HHI Measure	Point Estimate	95% CI
Suppliers		
Parent Group	+1	Superconsistent
Continent	1.03	(0.045,13.7)
Brand	24.2	(1.09,235)
Model	388	(363,898)
Competitive Advantage	-0.261	(-30.0,32.2)
Assemblers		
Parent Group	-0.0101	(-1.50,0.224)
Brand	-0.00789	(-2.07,0.831)
Model	-0.00437	(-3.64,34.2)
# Inequalities	298,272	
% Satisfied	82.3%	

The parameter on parent group specialization is fixed at +1. Estimating it with a smaller number of inequalities always finds the point estimate of +1, instead of −1. The estimate of a parameter that can take only two values is superconsistent, so I do not report a confidence interval. See Appendix B for details on estimation and inference.