Identification in Differentiated Products Markets Using Market Level Data*

Steven T. Berry
Yale University
Department of Economics
Cowles Foundation
and NBER

Philip A. Haile
Yale University
Department of Economics
Cowles Foundation
and NBER

September 23, 2009

Preliminary

Abstract

We consider nonparametric identification in models of differentiated products markets, using only market level observables. On the demand side we consider a nonparametric random utility model nesting random coefficients discrete choice models widely used in applied work. We allow for product/market-specific unobservables, endogenous product characteristics (e.g., prices), and high-dimensional taste shocks with arbitrary correlation and heteroskedasticity. On the supply side we specify marginal costs nonparametrically, allow for unobserved firm heterogeneity, and nest a variety of equilibrium oligopoly models. We pursue two approaches to identification. One relies on instrumental variables conditions used previously to demonstrate identification in a nonparametric regression framework. With this approach we can show identification of the demand side without reference to a particular supply model. Adding the supply side allows us to identify firms' marginal costs as well. Our second approach, more closely linked to classical identification arguments for supply and demand models, employs a change of variables approach. This leads to constructive identification results relying on exclusion and support conditions. We also point to testable implications that can discriminate between alternative models of oligopoly competition.

*Earlier drafts were circulated under the title "Identification of Discrete Choice Demand From Market Level Data." We had helpful early conversations on this topic with Rosa Matzkin and Yuichi Kitamura. We also had valuable feedback from Liran Einav, Jon Levin, Tim Bresnahan, and participants in the 2008 Game Theory World Congress, the 2008 LAMES, the 2009 FESAMES, 2009 SETA Conference, the Brown Conference on the Econometrics of Demand, and seminars at CREST, Penn, Stanford, USC, and UCSD. Financial support from the NSF is gratefully acknowledged.

1 Introduction

Discrete choice models of product differentiation play a central role in the modern empirical literature in industrial organization (IO) and are widely used in a range of applied fields of economics. Applications include studies of the sources of market power (e.g., Berry, Levinsohn, and Pakes (1995), Nevo (2001)), welfare gains from new goods or technologies (e.g., Petrin (2002), Eizenberg (2008)), mergers (e.g., Nevo (2000), Capps, Dranove, and Satterthwaite (2003)), network effects (e.g., Rysman (2004), Nair, Chintagunta, and Dube (2004)), product promotions (e.g., Chintagunta and Honoré (1996), Allenby and Rossi (1999)), environmental policy (e.g., Goldberg (1998)), vertical contracting (e.g., Villas-Boas (2007), Ho (2007)), market structure and product quality (e.g., Fan (2008)), media bias (e.g., Gentzkow and Shapiro (2009)), asymmetric information and insurance (e.g., Cardon and Hendel (2001), Bundorf, Levin, and Mahoney (2008), Lustig (2008)), trade policy (e.g., Goldberg (1995), Berry, Levinsohn, and Pakes (1999), Goldberg and Verboven (2001)), residential sorting (e.g., Bayer, Ferreira, and McMillan (2007)), and school choice (e.g., Hastings, Staiger, and Kane (2007)).

This work typically employs random utility discrete choice models allowing for rich heterogeneity in preferences, product/market-level unobservables, and endogenous prices Often the demand model is combined with a supply side involving an equilibrium oligopoly model of multi-product firms, allowing unobserved heterogeneity.

In practice these models are estimated using econometric specifications incorporating functional form restrictions and parametric distributional assumptions. Such restrictions may be desirable in practice: estimation in finite samples always requires approximations and, since the early work of McFadden (1974), an extensive literature has developed providing flexible discrete-choice models well suited to estimation and inference. Furthermore, parametric structure is necessary for the extrapolation involved in many out-of-sample counterfactuals. However, an important question is whether parametric specifications and distributional assumptions play a more fundamental role in determining what is learned from the data. In particular, are such assumptions essential for identification?

Here we explore nonparametric identification of this type of model, focusing on the common situation in which market level data are available, as in Berry, Levinsohn, and Pakes (1995) (henceforth, "BLP"). In such a setting, one observes market shares, market characteristics, product characteristics, and product/market level cost shifters, but not individual choices.[1] As usual, we do not assume observability of firms' marginal costs.

On the demand side, we consider a nonparametric generalization of random coefficients discrete choice models widely used in practice. We rely on an important restriction on the way product/market-specific unobservables enter preferences, but otherwise consider a very general random utility specification allowing for market/choice-specific unobservables, endogenous choice characteristics, unknown heteroskedasticity, and high-dimensional correlated taste shocks. We investigate identification of market-level demand as well as identification of the full random utility model. Identification of demand naturally requires instruments for prices (and/or any other endogenous choice characteristics), and we provide one result relying on standard nonparametric instrumental variables conditions (Newey and Powell (2003)). This demonstrates that the essential requirement for identification of demand in this type of model is the same as that for regression models: the availability of instruments. Further, this result can be extended to full identification of the random utility model by adding the same kind of separability and support conditions used to demonstrate identifiability of even the simplest semiparametric discrete response models.

Given identification of demand, we show that a generalization of standard parametric approaches to the estimation of oligopoly models leads us to nonparametric identification of firms' heterogeneous cost shocks and marginal cost functions, again relying on and index restriction and standard nonparametric instrumental variables conditions. This nonparametric identification of an oligopoly "supply side" provides a nonparametric foundation for many equilibrium counterfactual analyses performed in applied work.

While we view these as strong positive results, nonparametric instrumental variables conditions themselves can be difficult to interpret or verify. This is one reason we consider a

[1] We consider the case of "micro" (consumer-level) choice data in Berry and Haile (2009b).

second approach to identification, this time making simultaneous use of the demand model and a partially specified model of oligopoly competition. The supply side includes a set of generalized first-order conditions. Using the resulting system of "supply and demand" equations, we show that one can "invert" equilibrium conditions to obtain identification of demand and of an index that captures unobserved heterogeneity in marginal costs. If we further commit to a particular oligopoly model (e.g., Nash equilibrium in prices) we also recover firms' marginal costs.

Finally, we show that our model leads to testable restrictions that can discriminate between alternative models of oligopoly competition using either of our two identification approaches. This provides a generalization of important early results in the modern empirical IO literature on identifying the mode of competition (e.g., Bresnahan (1982, 1987)).

Together these results provide a positive message regarding the faith we may have in a growing body of applied work on differentiated products markets allowing for rich consumer and firm heterogeneity, choice-specific unobservables, and endogeneity. Such models are identified without parametric or distributional assumptions under the same sorts of conditions that identify simpler and more familiar models. Our results also shed light on the key sources of variation one should look for in applications.

To our knowledge, we provide the first and only results on the nonparametric identification of market-level discrete-choice differentiated products models of the sort found in BLP and other applications in IO. However, there is large related literature on the identification and estimation of semi- and nonparametric discrete choice models. On the demand side, our work is related to (and makes use of) much of this literature. Our work is also related to a large parametric literature on the estimation of "supply and demand" models, to a large literature on the estimation and testing of oligopoly models, and to work on the nonparametric identification of simultaneous equations models. In the following section we briefly place our work in the context of these and other prior literatures. We then set up the model in section 3 and discuss a key preliminary result in section 4. We provide our two sets of identification results in sections 5 and 6. Discrimination between alternative oligopoly models is discussed in 7. We conclude in section 8.

2 Related Literature

We examine identification of discrete-choice differentiated products models that are in the spirit of BLP and a large related applied literature. Two features are essential to these models: (i) rich heterogeneity in preferences and (ii) explicit modeling of product/market-level unobservables.[2] The former allows for flexible substitution patterns[3] (e.g., cross-price elasticities), while the latter is essential if one is to fully account for the endogeneity of prices implied by equilibrium pricing in oligopoly models.

This combination of factors has not been treated in the prior literature on identification. Indeed, although there is a large and closely related literature on identification of discrete-choice models,[4] there are no nonparametric or semiparametric identification results even for the linear random utility models used in this extensive applied literature.

Much of the literature considering the identification of discrete-choice models with heterogeneous preferences focuses on random coefficients models, but without endogeneity. Ichimura and Thompson (1998) studied a linear random coefficients binary choice model. Briesch, Chintagunta, and Matzkin (2005) consider multinomial choice, allowing some generalization of a linear random coefficients model. Our work relaxes functional form and distributional assumptions relied on in this earlier work, incorporates market/choice-specific unobservables, and allows for endogeneity through market/choice-specific unobservables.

[2]While our work is motivated by IO applications, these models are relevant in in many other discrete-choice contexts where there are unobservables at the level of a “group” (the analog of our “market”). For example, employees’ choices among offered insurance plans may depend on unobservable characteristics of the plans. A broad set of examples with “group level unobservables” is discussed in Berry and Haile (2009a) for the case of binary choice and related models. Although an oligopoly supply side may not be appropriate for all such examples, several results are obtained here without reference to a supply side, and the overall approach may be useful in other cases as well.

[3]See, e.g., the discussions in Domencich and McFadden (1975), Hausman and Wise (1978) and Berry, Levinsohn, and Pakes (1995). Early models of discrete choice with heterogeneous tastes for characteristics include those in Quandt (1966), Quandt (1968), and Domencich and McFadden (1975).

[4]Important early work includes Manski (1985), Manski (1988), Matzkin (1992), and Matzkin (1993), which examine semiparametric models with exogenous regressors.

A number of papers address the identification of discrete-choice models with endogeneity—sometimes only in a binary context, sometimes without consumer heterogeneity, and usually without the kind of endogeneity considered in the applied literature that motivates our work. Examples include Lewbel (2000), Honoré and Lewbel (2002), Hong and Tamer (2004), Blundell and Powell (2004), Lewbel (2005), and Magnac and Maurin (2007). These all consider linear semiparametric models, allowing for a single additive scalar shock (analogous to the extreme value or normal shock in logit and probit models) that may be correlated with some observables. Among these, Lewbel (2000) and Lewbel (2005) consider multinomial choice. Extensions to non-additive shocks are considered in Matzkin (2007a) and Matzkin (2007b).

Compared to these papers, we relax functional form restrictions and, more fundamental, add the important distinction between market/choice-specific unobservables and individual heterogeneity in preferences. This distinction allows the model to define comparative statics (e.g., responses of market shares to exogenous changes in price) that account for both heteroskedasticity (e.g., heterogeneity across in consumers' marginal rates of substitution between income and other characteristics) and endogeneity (e.g., correlation between a good's price and its unobserved quality).[5] For example, defining demand elasticities requires characterizing changes in market shares in response to a change in price, allowing the price change to affect the variance (and other moments) of utilities but holding fixed the market/choice-specific unobservables. Such comparative statics are essential to empirical work in the literatures that motivate our analysis, but cannot be defined in models with a single unobservable shock for each product.

Blundell and Powell (2004), Matzkin (2004), and Hoderlein (2008) have considered binary choice with endogeneity in semiparametric triangular models, leading to the applicability of control function methods or the related idea of "unobserved instruments" (see also Petrin and Train (2009), Altonji and Matzkin (2005), Gautier and Kitamura (2007), and Fox and Gandhi (2009)). However, standard models of oligopoly pricing in differentiated products

[5] Matzkin (2004) (section 5.1) makes a distinction between choice-specific unobservables and an additive preference shock, but in a model without random coefficients or other sources of heteroskedasticity/heterogeneous tastes for product characteristics. See also Matzkin (2007a) and Matzkin (2007b).

markets imply that each equilibrium price depends on the entire vector of demand and cost shocks. This rules out a triangular structure. Nonetheless, our "change of variables" approach uses a related strategy of inverting a multiproduct supply and demand system to recover the entire vector of shocks to costs and demand. This can be interpreted as a generalization of the control function approach.

On the supply side, Rosse (1970) introduced the idea of using first-order conditions for imperfectly competitive firms to infer their marginal costs from prices and demand parameters.[6] Our approach to identification of the oligopoly supply model is a nonparametric extension of that idea. Our insights regarding discrimination between alternative oligopoly models are closely related to the ideas in of Bresnahan (1982) and Lau (1982) (see also Bresnahan (1989) and references therein). They restricted attention to the class of conjectural variations models without structural errors, but a generalization of their idea of "rotations of demand" reappears in our framework. An insight from our results is that demand itself need not change to identify the correct oligopoly model: any change in the market that causes a firm's "residual marginal revenue function" to rotate can suffice.

Our change-of-variables approach, which exploits the simultaneous determination of prices and market shares, has links to the prior literature on the nonparametric identification of simultaneous equations models (e.g., Brown (1983), Roehrig (1988), Matzkin (2005), and Matzkin (2008)). A standard strategy in this literature is to relate the joint density of latent variables to that of the observables using restrictions from theory and a standard change of variables. A complication, emphasized in Benkard and Berry (2006), is that the change of variables involves the Jacobian of the transformation. This introduces substantial challenges and has limited the set of models for which identifiability has been shown using the change of variables approach. However, in our context the same index restriction that enables us to use the nonparametric instrumental variables approach enables us to use a new change of variables argument to obtain a constructive proof of identification. Here our work is

[6] Bresnahan (1989) provides a review of the early literature on oligopoly estimation that followed. BLP inverted the multiproduct oligopoly first-order conditions to solve for unobserved shocks to marginal cost.

closely related to that of Matzkin (2005, 2008), who has explored identification in a variety of nonparametric simultaneous equations models. Although she does not explicitly address discrete choice models, for our change of variable argument we transform our model to a form equivalent to one she considers. This transformation maps our index restriction to a separability condition whose advantages she emphasizes in a variety of other contexts. Even starting from the transformed model, however, our assumptions and proof differ from hers in important ways.[7] Our formal results may therefore complement those in Matzkin (2008) for applications of simultaneous equations even outside the application to discrete choice.

For a critical preliminary result, we rely heavily on insights in Gandhi (2008), which recently showed how to extend a key invertibility result of Berry (1994) and Berry and Pakes (2007) to a more general class of discrete choice demand models. We reinterpret Ghandi's key assumption graphically as an intuitive "connected substitutes" condition on the elements of the choice set. Although Gandhi (2008) focused on invertibility of demand, we show that the same demand-side assumption plays an important role in ensuring invertibility of the oligopoly supply side.

Turning to other recent unpublished papers, Berry and Haile (2009b) explores the identification of discrete choice models in the case of "micro data" relying in part on ideas similar to those used here. The distinction between "market data" and "micro data" has been emphasized in the recent industrial organization literature (e.g., Berry, Levinsohn, and Pakes (2004)), but not the econometrics literature on identification. A key insight in Berry and Haile (2009b) is that within a market all market/choice-specific unobservables are held fixed. One can therefore learn a great deal about the distribution of utilities from "variation in choice sets" created by within-market heterogeneity in consumer/choice-specific covariates—variation that is not confounded by variation in the market/choice-specific unobservables. That strategy is exploited throughout Berry and Haile (2009b), but cannot be applied to market level data.[8] In Berry and Haile (2009a) we have explored related ideas

[7]For example, we use the same independence and support assumptions she uses in discussing supply and demand, but we do not require any conditions on (even existence of) derivatives of densities.

[8]Berry and Haile (2009b) includes an example in which what appears to be a "market data" environment

in the context of a "generalized regression framework" (Han (1987)), which nests the binary choice model. For that class of models, the index restriction we require throughout the present paper can be dropped. Unfortunately, many applications fall outside the binary choice setting.

Concurrent work by Fox and Gandhi (2009) explores identifiability of several related models, including a flexible model of polychotomous choice in which consumer types are themselves multinomial and the conditional indirect utility functions are analytic. They do not consider our case with market-level data and endogenous prices set in a (non-triangular) system of equations.[9] Concurrent work by Chiappori and Komunjer (2009) considers a related change of variables approach in a "micro data" context.

3 Demand Model

3.1 Consumers, Products and Markets

Each consumer i in market t chooses from a set $\mathcal{J}_t$ of available products. We use the terms "good," "product," and "choice" interchangeably. The term "market" is synonymous with the choice set. In practice, markets will typically be defined geographically and/or temporally. The choice set always includes the option not to purchase, i.e., to choose the "outside good," which we index as choice $j = 0$. We denote the number of "inside goods" by $J_t = |\mathcal{J}_t| - 1$.[10]

Each inside good/market has observable (to us) characteristics $x_{jt} \in \mathbb{R}^{K_x}$ and price $p_{jt} \in \mathbb{R}$. We treat x_{jt} and p_{jt} differently because we will allow p_{jt} to be endogenous. The restriction

is actually isomorphic to the "micro data" environment. In that example one has a continuum of observations for which choice-specific unobservables are held fixed while observables vary. In general this is not the case.

[9] A recent paper by Bajari, Fox, Kim, and Ryan (2009) considers identification in a linear random coefficients model without endogeneity, assuming that the distribution of an additive i.i.d. preference shock is known.

[10] In applications with no "outside choice" our approach can be adapted by normalizing preferences relative to those for a given choice. The same adjustment applies when characteristics of the outside good vary across markets in observable ways.

to a single endogenous characteristic reflects the usual practice, but is not essential.[11] We allow x_{jt} to include product dummies. Likewise, x_{jt} can include components that vary only with the market t or only with the product j. Unobservables at the level of the product and market are represented by an index $\xi_{jt} \in \mathbb{R}$. In applications this is typically motivated by the presence of unobserved product characteristics and/or variation in unobserved tastes across markets. A market t is thus characterized by $\left(\mathcal{J}_t, \left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}\right)$. We let $\chi =$ supp$\left(x_{jt}, p_{jt}, \xi_{jt}\right)$ and $\chi^{\mathcal{J}_t} =$supp$\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}$.

3.2 Preferences

We consider preferences represented by a random utility model. Each consumer i in market t has conditional indirect utilities v_{ijt} for each product j determined by a function $u_{it} : \chi \to \mathbb{R}$. Consumers have heterogeneous tastes, even conditional on all observables. This is modeled by specifying each consumer i's utility function u_{it} as a random draw from a set $\mathcal{U}$. We discuss restrictions on $\mathcal{U}$ below.

More formally, let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space. Given any $\left(x_{jt}, p_{jt}, \xi_{jt}\right) \in \chi$, consumer i's conditional indirect utility from good j is given by

$$v_{ijt} = u_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right) = u\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right) \tag{1}$$

where u is measurable in ω_{it}, and $u\left(\cdot,\cdot,\cdot,\omega\right) \in \mathcal{U}$ for all $\omega \in \Omega$. Without loss, each consumer's draw ω_{it} from Ω is independent of $\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}$.[12]

This formulation superficially resembles models in which randomness in utilities is captured by a scalar random variable (e.g., Lewbel (2000), Matzkin (2007a), Matzkin (2007b)); however, here ω_{it} is not a random variable but an elementary event in Ω that can determine an arbitrary number of random variables. The following example illustrates.

[11] The modifications required to allow higher dimensional p_{jt} are straightforward, although the usual challenge of finding adequate instruments for more than one endogenous product characteristic would remain.

[12] This is standard for defining a random function. It is without loss because the function u already allows the distribution of v_{ijt} to depend freely on x_{jt}, p_{jt}, ξ_{jt}.

Example 1. *A special case of the class of preferences we consider is generated by the linear random coefficients random utility model*

$$u\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right) = x_{jt}\beta_{it} - \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{2}$$

where $(\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ *are defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *as*

$$\left(\alpha\left(\omega_{it}\right), \beta^{(1)}\left(\omega_{it}\right), \ldots, \beta^{(K_x)}\left(\omega_{it}\right), \epsilon_1\left(\omega_{it}\right), \ldots, \epsilon_J\left(\omega_{it}\right)\right).$$ [13]

This structure permits $(\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ *to have an arbitrary joint distribution but could be relaxed further. For example specifying* $\epsilon_{ijt} = \epsilon_j\left(x_{jt}, p_{jt}, \omega_{it}\right)$ *would allow richer preference heterogeneity. If we relax the model further by specifying* $\epsilon_{ijt} = \epsilon_j\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right)$*, the terms* $x_{jt}\beta_{it} - \alpha_{it}p_{jt} + \xi_{jt}$ *in (2) become redundant and we obtain our original model* $u\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right) = \epsilon\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right)$*.*

Note that there is no market subscript t on the probability measure $\mathbb{P}$. This reflects our assumption that the ξ_{jt} captures all unobserved heterogeneity at the market and/or product level. This is standard in the literature but is an important restriction.[14] Aside from this restriction, however, our representation of preferences is so far fully general. For example, it allows arbitrary correlation of consumer-specific tastes for different goods or characteristics. It also allows arbitrary heteroskedasticity in utilities across different products, or in utilities for a given product as its characteristics $\left(x_{jt}, \xi_{jt}\right)$ vary.

We will, however, rely on the following important restriction on preferences

---

[13] The fact that we allow product dummies as components of x_{jt} enables us to write choice-specific functions like ϵ_j here. Note also that this structure permits variation in J_t across markets. The realization of ω_{it} should be thought of as generating values of $\epsilon_{ijt} = \epsilon_j\left(\omega_{it}\right)$ for all possible choices j, not just those in the current choice set. Thus, the utility function defines preferences even over products not available.

[14] An exception is Athey and Imbens (2007), although they do not address identifiability of their model. Athey and Imbens (2007) and Berry and Haile (2009b) point out testable restrictions in "micro data" settings if one assumes that the same scalar product/market-level unobservable applies to multiple subpopulations of consumers. The model of binary choice nested in the generalized regression model of Berry and Haile (2009a) permits a different unobservable for every vector of consumer-level observables.

Assumption 1a. $\mathcal{U}$ is the set of all functions $\tilde{u}_{it} : \chi \to \mathbb{R}$ such that, (i) for any $\mathcal{J}_t$, any distinct $k, \ell \in \mathcal{J}_t$, and any $\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \{k,\ell\}}$, the random differences $\tilde{u}_{it}\left(x_{kt}, p_{kt}, \xi_{kt}\right) - \tilde{u}_{it}\left(x_{\ell t}, p_{,\ell t}, \xi_{\ell t}\right)$ are continuously distributed with convex support; (ii) $\tilde{u}_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right) = \mu_{it}\left(x_{jt}^{(1)} + \xi_{jt}, x_{jt}^{(2)}, p_{jt}\right)$ for some function μ_{it} that is strictly increasing in its first argument.

Part (i) simplifies the analysis by enabling us to ignore both ties and choice probabilities that are invariant to a strict (stochastic) increase in v_{ijt} for some j. Part (ii) is the more substantive restriction and limits attention to random utility functions admitting representations of the form

$$v_{ijt} = u\left(x_{jt}^{(1)} + \xi_{jt}, x_{jt}^{(2)}, p_{jt}, \omega_{it}\right). \tag{3}$$

There are two parts to this restriction. The first is a linear index restriction, requiring perfect substitutability between ξ_{jt} and $x_{jt}^{(1)}$ inside the function u. As mentioned already, this plays a key role in both of our identification approaches. The second is a restriction to a "vertical" unobservable ξ_{jt}: all else equal, an increase in ξ_{jt} makes product j more attractive to all consumers in market t. This is standard in the applied literature but is a restriction we rely on to allow recovery of ξ_{jt} for every product.[15] In the the linear random coefficients model of Example 1, Assumption 1a holds if one covariate enters with a fixed coefficient.[16]

With the structure on preferences in Assumption 1a we will be able to show identification of demand, identification of marginal costs, and the testability of a given oligopoly model. For our results on full identification of the random utility model (defined below) we will rely on a more restrictive specification.

Assumption 1b. $\mathcal{U}$ is the set of all functions $\tilde{u}_{it} : \chi \to \mathbb{R}$ such that (i) for any $\mathcal{J}_t$, any distinct $k, \ell \in \mathcal{J}_t$, and any $\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \{k,\ell\}}$, the random differences $\tilde{u}_{it}\left(x_{kt}, p_{kt}, \xi_{kt}\right) -$

[15] Berry and Haile (2009b) discuss testable implications in the case of micro data.

[16] To relate this to models like those in Ichimura and Thompson (1998), Lewbel (2000), Briesch, Chintagunta, and Matzkin (2005), or Gautier and Kitamura (2007) that exclude the market/choice-specific unobservable ξ_{jt}, note that without ξ_{jt} part (ii) of Assumption 1a is satisfied as long as v_{ijt} is strictly increasing in some $x_{jt}^{(k)} \in x_{jt}$ with probability one.

$\tilde{u}_{it}(x_{\ell t}, p_{,\ell t}, \xi_{\ell t})$ are continuously distributed with convex support; (ii) for all i and t there is a monotonic function Γ_{it} such that $\Gamma_{it}(\tilde{u}_{it}(x_{jt}, p_{jt}, \xi_{jt})) = x_{jt}^{(1)} + \xi_{jt} + \mu_{it}\left(x_{jt}^{(2)}, p_{jt}\right)$ for some function μ_{it}.

This assumption differs from Assumption 1a in requiring conditional indirect utilities with representations of the form

$$v_{ijt} = x_{jt}^{(1)} + \xi_{jt} + \mu\left(x_{jt}^{(2)}, p_{jt}, \omega_{it}\right). \tag{4}$$

As is standard, we rely on additive separability in an exogenous observable to provide a mapping between the observe units of probability and the latent units of utility. In addition, (4) provides a specification of cardinal utility that, under standard conditions, could be used to characterize a variety of welfare measures (in aggregate or across subpopulations defined by observables).[17]

Finally, two types of normalizations will be needed to obtain a unique representation of preferences. Such normalizations are without loss of generality. One is a normalization of utilities, which have no natural location or units (scale). Throughout the paper we normalize the location of utilities by setting the utility from the outside good to zero: $v_{i0t} = 0$. We will require a scale normalization of utilities only for the results using Assumption 1b, which already normalizes the scale of utility. The second type of normalization involves the choice-specific unobservables ξ_{jt}. Linear substitutability between $x_{jt}^{(1)}$ and ξ_{jt} (Assumption 1a) already defines the scale of each ξ_{jt}. Because it will be convenient to use a different location normalization of ξ_{jt} for each of our two identification approaches, we will provide these below.

[17] For example, this model could be used to define aggregate compensating/equivalent variation in units of the normalized marginal utility for the numeraire characteristic $x_{jt}^{(1)}$. This typically will not allow characterization of Pareto improvements, since marginal utility from $x_{jt}^{(1)}$ will not be transferable. Adding an assumption of quasilinearity in wealth (price) would enable calculations in units of money.

3.3 Market Shares

Given the choice set, each consumer maximizes her utility, choosing product j whenever

$$u\left(x_{jt}^{(1)}+\xi_{jt},x_{jt}^{(2)},p_{jt},\omega_{it}\right)>u\left(x_{kt}^{(1)}+\xi_{kt},x_{jt}^{(2)},p_{kt},\omega_{it}\right)\quad\forall k\in\mathcal{J}_t-\{j\}. \tag{5}$$

We denote consumer i's choice by

$$y_{it}=\arg\max_{j\in\mathcal{J}_t}\ u\left(x_{jt},\xi_{jt},z_{ijt},\omega_{it}\right)$$

and market shares (choice probabilities) by

$$\begin{aligned}
s_{jt} &= E_{\mathbb{P}}\left[1\left\{y_{it}=j\right\}\right]\\
&= E_{\mathbb{P}}\left[1\left\{u\left(x_{jt},p_{jt},\xi_{jt},\omega_{it}\right)>u\left(x_{kt},p_{kt},\xi_{kt},\omega_{it}\right)\ \forall k\in\mathcal{J}_t-\{j\}\right\}\right]\\
&\equiv s_j\left(\mathcal{J}_t,\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}_t}\right). \qquad (6)
\end{aligned}$$

We will assume that all goods observed in equilibrium have positive market shares.

Assumption 2. For all $\mathcal{J}_t$ and $\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}_t}\in\chi^{\mathcal{J}_t}$, $s_j\left(\mathcal{J}_t,\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}_t}\right)>0$ for all j with probability one.

Market shares are positive for all goods at *any* $\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}_t}$ in models for which the support of $\{v_{ijt}\}_{k\in\mathcal{J}_t}$ is always $\mathbb{R}^{J_t}$. In a parametric context this includes multinomial probit or logit models. Of course, Assumption 2 requires only that market shares be positive at *equilibrium* values of $\mathcal{J}_t,\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}_t}$.[18] This requirement is, of course, directly testable.

[18] For example, in the usual price-setting oligopoly model an inside good can have zero market share only if a good is offered by a firm even though it cannot be sold at any positive price. A necessary condition for the outside good to have zero market share is that the there be no downward distortion in market output due to imperfect competition. This can arise in some simple oligopoly models like that of Hotelling (1929) if preferences and locations are such that the market is "covered."

3.4 Observables and Primitives of Interest

We let M_t denote the measure of consumers in market t (the "market size"), and let $\tilde{z}_t$ denote instruments excludable from the utility function. The observables consist of $(t, M_t, \mathcal{J}_t, \tilde{z}_t, \{s_{jt}, x_{jt}, p_{jt}\}_{j\in\mathcal{J}_t})$. To discuss identification, for every $(t, M_t, \mathcal{J}_t)$ we treat the population distribution of $(\tilde{z}_t, \{s_{jt}, x_{jt}, p_{jt}\}_{j\in\mathcal{J}_t})$ as known.

On the demand side of the market, we will consider two types of identification results. One is *identification of demand*; i.e., identification of each ξ_{jt} and the functions s_j defined in (6). These primitives fully characterize the demand system: they describe how product characteristics (observed and unobserved, endogenous and exogenous) determine the market shares of all goods, including the outside good.

Identification of demand is sufficient for many purposes motivating demand estimation. This includes analysis of consumer surplus and the use of the demand structure with a supply model to infer markups. However, we also consider identification of the joint distribution of indirect utilities conditional on the choice set $(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$. These conditional distributions are the primitives determining all quantities defined by our model.[19] We refer to this as *full identification of the random utility model.*

Henceforth we will condition on $\mathcal{J}_t = \mathcal{J}$ with $|\mathcal{J}| = J$. We also condition on a value of $x_t^{(2)} = \left(x_{1t}^{(2)}, \ldots, x_{Jt}^{(2)}\right)$ and suppress $x_t^{(2)}$ in the notation. For simplicity we then let x_{jt} represent $x_{jt}^{(1)}$. Conditioning on $x_t^{(2)}$requires that rewrite

$$v_{ijt} = u_j\left(x_{jt} + \xi_{jt}, p_{jt}, \omega_{it}\right) \tag{7}$$

and

$$v_{ijt} = x_{jt} + \xi_{jt} + \mu_j\left(p_{jt}, \omega_{it}\right) \tag{8}$$

[19] However, they do not define individual utilities. It should not be surprising the market level observables cannot be used to identify Pareto improvements. But in fact, only functional form restrictions that link each consumer to a realization of a random parameter can enable a random utility model to define changes in individual utilities, even in a micro data setting. We discuss this issue further in Berry and Haile (2009b), where we also point to conditions that can be used to extend identification of the conditional joint distribution of utilities to to identification of linear random coefficients models, which do impose structure sufficient to define changes in individual utilities.

to represent, respectively, (3) and (4) above, since the utility functions will generally be evaluated at different $x_{jt}^{(2)}$ for each j. We will work with these two representations of preferences in what follows.

4 Connected Substitutes

Central to our identification arguments is the inversion of equilibrium conditions—of choice probabilities implied by utility maximization on the demand side and of first-order conditions for oligopoly equilibrium the supply side. A key condition that the choice set $\mathcal{J}$ be comprised of substitute goods that "belong" (in a sense defined below) in the same choice problem for at least some consumers. To state this "connected substitutes" assumption, we first need a definition.

Definition 1. *Product k substitutes to product ℓ at $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}$* if, for any variable w_{jt} such that v_{ijt} is strictly stochastically increasing in w_{jt}, $s_\ell\left(\mathcal{J}, \{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}\right)$ rises when w_{kt} falls.[20]

Our initial use of this definition involves the case $w_{kt} = x_{kt} + \xi_{kt}$, although we also consider $w_{kt} = -p_{kt}$ when we discuss identification of supply. Definition 1 provides a directional notion of one product being a substitute for another. For example, if v_{ijt} is strictly decreasing in p_{jt}, we would say that product k substitutes to product ℓ if a rise in p_{kt} leads (all else equal) to a larger market share for product ℓ.

Given a value of $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}$, let $\Sigma(\mathcal{J})$ denote the $(J+1)\times(J+1)$ matrix of zeros and ones, with the (r,c) element equal to one if product $(r-1)$ substitutes to product $(c-1)$.

Assumption 3 ("Connected Substitutes"). For all $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}} \in \chi^{\mathcal{J}}$, the directed graph of $\Sigma(\mathcal{J})$ is strongly connected.

---

[20] Because we introduce this assumption after normalizing the utility of the outside good to zero, we define a fall in w_{0t} to mean equal increases in w_{jt} for all $j > 0$. Thus product 0 substitutes to product j if the market share of product j increases whenever w_{kt} increases by an equal amount for all $k > 0$.

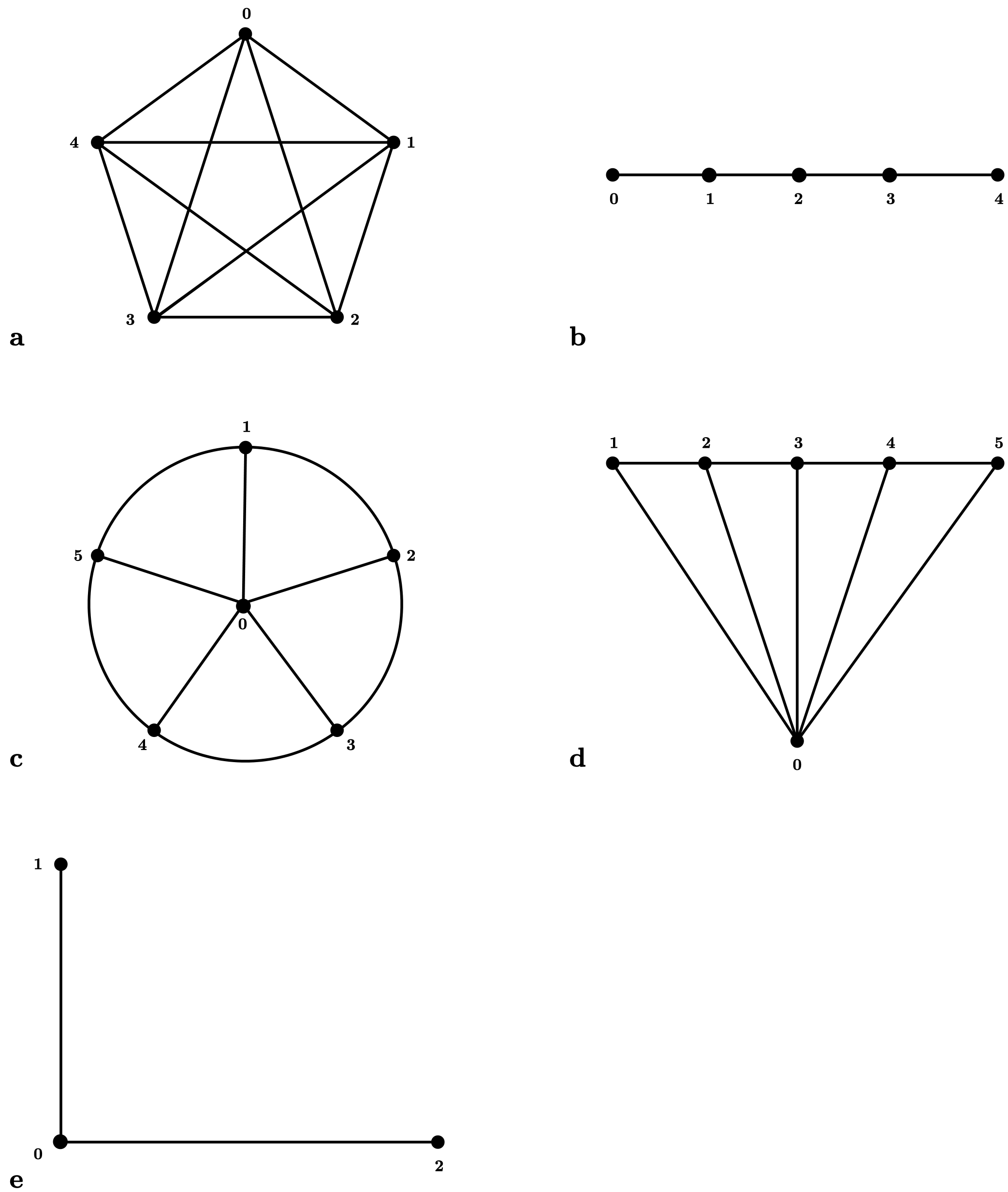

Figure 1: Graphs of $\Sigma(\mathcal{J})$ at generic $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j \in \mathcal{J}}$ for standard models. All edges are bi-directional, so for simplicity we show undirected graphs. Panel a: multinomial logit, multinomial probit, mixed logit, etc.; Panel b: pure vertical models e.g., (, Mussa and Rosen (1978), Bresnahan (1981), etc.); Panel c: Salop (1979) (with an outside good); Panel d: Rochet and Stole (2002); Panel e: independent goods and an outside good.

The directed graph of $\Sigma(\mathcal{J})$ has nodes (vertices) representing each product and an edge from product k to product ℓ whenever product k substitutes to product ℓ.[21] The connected substitutes condition requires that for any distinct products j and j', there be a path of substitution, possibly indirect, from j to j'. If this condition fails, then there is some subset of products that substitute only among themselves for at least some values of $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}$. It is natural to rule this out,[22] and Assumption 3 is indeed satisfied in standard models. For example, in multinomial probit and logit models (with or without random coefficients), every product substitutes to every other product. Figure 1 illustrates the directed graphs of $\Sigma(\mathcal{J})$ for a variety of models. Panel e of this figure demonstrates that even a market comprised of two independent goods satisfies this condition as long as each substitutes to and from the outside good.

5 Identification with General IV Conditions

5.1 Identification of Demand

Let

$$\delta_{jt} = x_{jt} + \xi_{jt}.$$

Let $x_t = (x_{1t}, \ldots, x_{Jt})$, $p_t = (p_{1t}, \ldots, p_{Jt})$, and $\delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$.

The key implication of Assumption 1a is that choice probabilities depend on the indices δ_{jt} rather than separately on its components x_{jt} and ξ_{jt}. In particular, for any vector δ_t,

[21] In standard examples $\Sigma(\mathcal{J})$ is symmetric, so its directed graph can be represented with bi-directional edges. See the additional discussion in Appendix D.

[22] Gandhi (2008) used an equivalent assumption (with $w_{kt} = \xi_{kt}$) to show the invertibility of demand, a result we utilize below. We show that the same condition plays an important role in the invertibility of the supply side as well. See Lemma 3 in Appendix A for the equivalence of our connected substitutes assumption and Gandhi's condition.

market shares are given by

$$\begin{aligned} s_{jt} &= E_{\mathbb{P}}\left[1\left\{u_j\left(\delta_{jt}, p_{jt}, \omega_{it}\right) \geq u_k\left(\delta_{kt}, p_{jt}, \omega_{it}\right)\ \ \forall k \in \mathcal{J}\right\}\right] \\ &= \int_{v: v_j \geq v_k \forall k} dF_v\left(v | \delta_t, p_t\right) \\ &\equiv \sigma_j\left(\delta_t, p_t\right) \end{aligned} \tag{9}$$

where v is a J-vector and $F_v\left(v_{it}|\delta_t, p_t\right)$ is the joint distribution of $v_{it} = (v_{i1t}, \ldots, v_{iJt})$ conditional on (δ_t, p_t).

Using the connected substitutes assumption, we can follow the argument in Theorem 2 of Gandhi (2008) to show the following lemma, which generalizes well-known invertibility results for linear discrete choice models in Berry (1994) and Berry and Pakes (2007).[23]

Lemma 1. *Consider any price vector p and any market share vector $s = (s_1, \ldots, s_J)'$ on the interior of $\triangle^J$. Under Assumptions 1a and 3, there is at most one vector δ such that $\sigma_j\left(\delta, p\right) = s_j\ \forall\ j$.*

Proof. See Appendix B.

Using this result, we have

$$\delta_{jt} = \sigma_j^{-1}\left(s_t, p_t\right) \quad \forall j \tag{10}$$

which we can rewrite as

$$x_{jt} + \xi_{jt} = \sigma_j^{-1}\left(s_t, p_t\right) \quad \forall j. \tag{11}$$

To state the instrumental variables conditions, recall that $\tilde{z}_t$ represents instruments for p_t excluded from the determinants of $\{v_{ijt}\}_{j\in\mathcal{J}}$. Standard excluded instruments include cost shifters or prices of the same good in other markets (Hausman (1996), Nevo (2001)).

[23]Berry (1994) and Berry and Pakes (2007) show existence and uniqueness of an inverse choice probability in models with an additively separable δ_{jt}. Gandhi (2008) relaxes the separability requirement. Our lemma addresses only uniqueness conditional on existence. Under our maintained assumption that the model is correctly specified, given any observed choice probability vector, there must exist a vector $(\delta_1, \ldots, \delta_J)$ that rationalizes it. Gandhi (2008) provides conditions guaranteeing that an inverse exists for every choice probability vector in $\triangle^J$. Our uniqueness result differs from his only slightly, mainly in recognizing that the argument applies to a somewhat more general model of preferences.

With the exogenous condition variables $(\tilde{z}_t, x_t)$, we take the following instrumental variables assumptions from Newey and Powell (2003).

Assumption 4. $E[\xi_{jt}|\tilde{z}_t, x_t] = 0$ almost surely.

Assumption 5. For any function $B(s_t, p_t)$ with finite expectation, if $E[B(s_t, p_t)|\tilde{z}_t, x_t] = 0$ almost surely then $B(s_t, p_t) = 0$ almost surely.

Assumption 4 is a standard exclusion restriction, requiring mean independence between the instruments and the structural error ξ_{jt}. Note that setting $E[\xi_{jt}|\tilde{z}_t, x_t]$ equal to zero rather than another constant provides the required normalization of the location of ξ_{jt}. Assumption 5 is a "completeness" condition, which is the nonparametric analog of the standard rank condition for linear models. This condition requires that the instruments move the endogenous variables (s_t, p_t) sufficiently to ensure that any function of these variables can be distinguished from other functions through the exogenous variation in the instruments. Lehman and Romano (2005) gives standard sufficient conditions for completeness, and additional discussion is provided in Newey and Powell (2003) and Severini and Tripathi (2006).

Newey and Powell (2003) used analogs of Assumptions 4 and 5 to show the identifiability of a separable nonparametric regression model. The following result shows that the same argument can be applied to show identification of demand in our discrete choice setting.[24]

Theorem 1. *Under Assumptions 1a and 2–5, for all j (i) ξ_{jt} is identified for all t, and (ii) the function $s_j(\mathcal{J}, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}})$ is identified on $\chi^{\mathcal{J}}$.*

Proof. For any j, rewriting (11) and taking expectations conditional on z_{jt}, we obtain

$$E\left[\xi_{jt}|\tilde{z}_t, x_t\right] = E\left[\sigma_j^{-1}(s_t, p_t)\middle|\, \tilde{z}_t, x_t\right] - x_{jt}$$

[24] If we assumed a bounded support for ξ_{jt} and x_{jt} we could replace the completeness condition with bounded completeness to show identification of demand.

so that by Assumption 4,

$$E\left[\sigma_j^{-1}\left(s_t,p_t\right)\middle|\,\tilde{z}_t,x_t\right]-x_{jt}=0 \qquad a.s.$$

Suppose there is another function $\tilde{\sigma}_j^{-1}$ satisfying

$$E\left[\tilde{\sigma}_j^{-1}\left(s_t,p_t\right)\middle|\,\tilde{z}_t,x_t\right]-x_{jt}=0 \qquad a.s.$$

Letting $B\left(s_t,p_t\right)=\sigma_j^{-1}\left(s_t,p_t\right)-\tilde{\sigma}_j^{-1}\left(s_t,p_t\right)$, this implies

$$E\left[B\left(s_t,p_t\right)|\tilde{z}_t,x_t\right]=0 \qquad a.s.$$

But by Assumption 5 this requires $\tilde{\sigma}_j^{-1}=\sigma_j^{-1}$ almost surely, implying that σ_j^{-1} is identified. Repeating for all j, each ξ_{jt} is then uniquely determined by (11), proving part (i). Because choice probabilities are observed and all arguments of the demand functions $s_j\left(\mathcal{J},\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}}\right)$ are now known, part (ii) follows immediately. □

The proof is very similar to that given by Newey and Powell (2003) in the context of nonparametric regression. A difference is that we have, in addition to the nonparametric function $\sigma_j^{-1}\left(s_t,p_t\right)$, the additive x_{jt}. This drops out in the proof, which is important because it allows use to use x_{jt} as an instrument. This is in fact the reason we require that $x_{jt}+\xi_{jt}$ enter utility as an index.

5.2 Full Identification of the Random Utility Model

To obtain full identification, we rely on the quasilinear model of preferences and add a large support assumption:

Assumption 6. supp $x_t=\mathbb{R}^J$.

This support condition is strong. However, it is also standard in the literature because it provides a natural benchmark for evaluating identification under ideal conditions on observables. It is intuitive that in order to trace out the full CDF of the random part of a random utility model, extreme values of observables will be needed.[25] As the proof of the following result makes clear, we use the large support condition only for this role; in particular, we do not use the common "identification at infinity" argument that takes observables for all but one choice to extreme values in order to reduce a multinomial choice problem to a binary choice problem.

Theorem 2. *Under Assumptions 1b and 2–6, the joint distribution of* $(v_{i1t}, \ldots, v_{iJt})$ *conditional on any* $\{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}} \in \chi^{\mathcal{J}}$ *is identified.*

Proof. The market share of the outside good, conditional on p_t, x_t, and $(\xi_{1t}, \ldots, \xi_{Jt})$, is

$$Pr\left(\mu_1\left(p_{1t}, \omega_{it}\right) \leq -x_{1t} - \xi_{1t}, \ldots, \mu_J\left(p_{Jt}, \omega_{it}\right) \leq -x_{Jt} - \xi_{Jt}\right).$$

By Theorem 1 each ξ_{jt} is identified. Thus, under Assumption 6, variation in the vector x_t identifies the entire joint distribution of

$$\left(\mu_1\left(p_{1t}, \omega_{it}\right), \ldots, \mu_J\left(p_{Jt}, \omega_{it}\right)\right)$$

for any $\{p_{kt}\}_{k \in \mathcal{J}}$ in their support. Identification of the joint distribution of utilities conditional on any $\{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}} \in \chi^{\mathcal{J}}$ then follows. □

5.3 Adding A Supply Side

If we are willing to add a specification of the supply model, we can obtain identification of firms' marginal costs as well, although here we can return to the less restrictive specification of preferences in (3). Our approach generalizes arguments from the parametric literature on the estimation of static oligopoly models, which utilize first-order conditions for firms

[25] To our knowledge, all results showing semiparametric or nonparametric identification of a full random utility model rely on a similar condition (e.g., Matzkin (1992), Matzkin (1993), Ichimura and Thompson (1998), Lewbel (2000), Fox and Gandhi (2009)).

to solve for marginal costs in terms of demand parameters. Use of first-order conditions requires that the market share function $\sigma_j(\delta_t, p_t)$ be differentiable with respect to prices. It is easy to confirm that this holds in standard models, and we will assume it directly.

Assumption 7. $\sigma_k(\delta_t, p_t)$ is continuously differentiable with respect to p_ℓ $\forall k, \ell \in \mathcal{J}$.

We consider a nonparametric specification of costs, but require sufficient structure to ensure that behavior be characterized by first-order conditions that can be inverted to solve for the latent cost shocks. As with the demand model, the most restrictive assumption we require is an index restriction on how unobservables enter. We assume the marginal cost associated with product j depends on its output quantity, a "cost index," and other cost shifters:

$$mc_j\left(q_{jt}, z_{jt}^{(1)} + \eta_{jt}, z_{jt}^{(2)}\right). \tag{12}$$

Here η_{jt} is an unobserved cost shock and $\left(z_{jt}^{(1)}, z_{jt}^{(2)}\right)$ are observed cost shifters, with $z_{jt}^{(1)} \in \mathbb{R}$. We permit $z_{jt}^{(2)}$ to include components of $x_{jt}^{(2)}$, although $x_{jt}^{(1)}$ is excluded. We will be explicit below about the independent variation required of $z_{1t}^{(1)}, \ldots, z_{Jt}^{(1)}$. Output quantity for product j is $q_{jt} = M_t\sigma_j(p_t, \delta_t)$. Parallel to our model of demand, (12) assumes perfect substitution between the unobserved cost shock η_{jt} and the cost shifter $z_{jt}^{(1)}$. This is an important restriction, but one that is satisfied in many standard models.[26] We denote the cost index by $\zeta_{jt} \equiv z_{jt}^{(1)} + \eta_{jt}$.

We continue to condition on (and suppress) $x_t^{(2)}$. We now also condition on a value of $z_t^{(2)} = \left(z_{1t}^{(2)}, \ldots, z_{Jt}^{(2)}\right)$, likewise suppressing it in the notation and letting z_{jt} denote $z_{jt}^{(1)}$ for simplicity. We will show invertibility of the supply model under the following conditions.

Assumption 8. For all j

(i) $mc_j(q_{jt}, \zeta_{jt})$ is strictly monotonic in ζ_{jt};

(ii) $u_j(\delta_{jt}, p_{jt}, \omega_{it})$ is strictly decreasing in p_{jt};

(iii) there exists a function ψ_j (possibly unknown) such that for any equilibrium value of

[26] Note that $z_{jt}^{(1)}$ and η_{jt} could be any known transformations of some other observed and unobserved cost shifters.

(s_t, p_t) with $s_{jt} > 0 \forall j$,

$$mc_j(M_t s_{jt}, \zeta_{jt}) = \psi_j\left(s_t, M_t, D_t\left(s_t, p_t\right), p_t\right)$$

where $D_t\left(s_t, p_t\right)$ is the $J \times J$ matrix of partial derivatives $\left\{\frac{\partial \sigma_k(p_t, \delta_t)}{\partial p_\ell}\right\}_{k,\ell}$.

Given the index restriction, part (i) of Assumption 8 is fairly weak: weak monotonicity in an unobservable could be assumed without loss, since that would merely define an order on the unobservable. Part (ii) requires strictly downward sloping demand. Part (iii) is a high-level condition and is more opaque. It requires that it be possible to rewrite first-order conditions to express marginal cost for each product as a function of equilibrium quantities (market shares), prices, and derivatives of demand at these prices and quantities. We show in Appendix C that this holds for a variety of standard oligopoly models, including the multi-product price-setting oligopoly model most often used in empirical work. By relying on Assumption 8 we provide results for a class of models rather than just one. This will be particularly useful when we discuss discrimination between alternative models.

The following lemma shows that, given the demand structure, Assumption 8 is sufficient to ensure that first-order conditions provide an invertible map from the cost indices ζ_t to equilibrium prices.

Lemma 2. *For any market size M_t, there is exactly one $(\zeta_{1t}, \ldots, \zeta_{Jt}) \in \mathbb{R}^J$ consistent with Assumption 8.*

Proof. By part (i) of Assumption 8, the function mc_j in part (iii) can be inverted, yielding

$$\zeta_{jt} = mc_j^{-1}\left(M_t s_{jt}, \psi_j(s_t, M_t, D_t\left(s_t, p_t\right)_t, p_t)\right). \tag{13}$$

Given M_t and the fact that $s_{jt} = \sigma_j\left(\delta_j, p_t\right)$, the right-hand side is a function of δ_t, p_t. □

Henceforth we will fix a value of M_t and suppress it in the notation. We can then re-write

(13) as[27]

$$z_{jt} + \eta_{jt} = \pi_j^{-1}(s_t, p_t) \quad \forall j. \tag{14}$$

This provides a key set of equations for what follows.

Note that equation (14) takes the same form as (11). We will use this relation in the same way. Let $\tilde{z}_t = (z_{1t}, \ldots, z_{Jt})$; i.e., the excluded instruments on the demand side are the exogenous cost shifters from the supply model. We can then show the following result.

Theorem 3. *Suppose that Assumptions 1a, 2–5, 7 and 8 hold. Then for all j (i) η_{jt} is identified for all t, and (ii) if each ψ_j is known, the function $mc_j(q_{jt}, \zeta_{jt})$ is identified on the support of (q_{jt}, ζ_{jt}).*

Proof. Part (i) follows by observing that the argument used in the proof of Theorem 1 can be repeated with trivial modification to recover the inverse pricing relations π_j^{-1} and the cost shocks ζ_{jt} using the instrumental variables $(x_t, \tilde{z}_t)$. Now observe that

$$mc_j(q_{jt}, \zeta_{jt}) = \psi_j(s_t, D_t(s_t, p_t), p_t). \tag{15}$$

Theorem 1 ensures that $D_t(s_t, p_t)$ is known. Thus all arguments of ψ_j are known. Since ψ_j is itself known, the right side of (15) is known. Since $q_{jt} = M_t s_{jt}$, by part (i) both arguments of the left side of (15) are known. Part (ii) then follows. □

6 A Change of Variables Approach

The preceding analysis yields encouraging results. A flexible model of demand (and supply) for differentiated products is identified under the same kind of instrumental variables conditions required for identification of regression models. Full identification holds as well if we add the kind of separability and support conditions used to show identification of even

[27] The function π_j^{-1} involves the composition of mc_j^{-1} and ψ_j. Although we do not define a function π_j, we use the notation π_j^{-1} as a reminder that this represents an "inversion" of supply side equilibrium conditions.

the simplest semiparametric models of multinomial choice. However, a limitation of the results above is the abstract nature of the completeness condition, which can be difficult to interpret or verify. Here we consider an alternative approach that treats the demand and supply models as a system. This enables us to pursue a change of variables argument often useful for simultaneous equations models (e.g., Brown (1983), Roehrig (1988), Matzkin (2005), and Matzkin (2008)).

This approach has advantages and disadvantages relative to the previous approach. The main disadvantages are the need to place some structure on the supply side even to identify demand, and the need for additional conditions ensuring that we can relate a joint density of the latent structural errors to a joint density of observables. These involve regularity conditions as well as a high level assumption to avoid problems that can be created by multiple equilibria. An advantage is that we will be able to replace the abstract completeness condition with a support condition on demand and cost shifters. This leads to constructive arguments with close connections to classical identification arguments for models of demand and supply.

We begin with results on identification of demand, which we can show without fully specifying the supply side. We then address identification of marginal costs under a complete specification of the supply model.

6.1 Identification of Demand and the Random Utility Model

Because we treat demand and supply as a system here, the change of variables approach requires that the map (14) between (s_t, p_t) and the cost indices ζ_t be one-to-one. We will assume this directly.

Assumption 9. There is a unique vector of equilibrium prices associated with any (δ, ζ).

This assumption is satisfied if the equilibrium first-order conditions have a unique solution *in price* for any (δ, ζ), conditional on given marginal cost and demand functions. This is often difficult to verify in models of product differentiation (see, for example, Caplin and

Nalebuff (1991)), and it is not hard to construct examples in which multiple equilibria do exist. If there are multiple equilibria, Assumption 9 requires an equilibrium selection rule such that the same prices p_t arise whenever (δ_t, ζ_t) is the same. This this rules out random equilibrium selection or equilibrium selection based on x_{jt} or ξ_{jt} instead of their sum δ_{jt} (and similarly for ζ_{jt}).

From the analysis above we repeat the two equations (11) and (14):

$$\begin{aligned} x_{jt} + \xi_{jt} &= \sigma_j^{-1}(s_t, p_t) \quad \forall j \\ z_{jt} + \eta_{jt} &= \pi_j^{-1}(s_t, p_t) \quad \forall j. \end{aligned}$$

We will considered these $2J$ equations as a system.

As in the preceding sections, the linear structure of the indices normalizes the scale of the unobservables ξ_{jt} and η_{jt}. To normalize locations, instead of setting means to zero, without loss we take any (x^0, z^0) and any (s^0, p^0) in the support of $(s_t, p_t) \,|\, (x^0, z^0)$ and let

$$\begin{aligned} \sigma_j^{-1}\left(s^0, p^0\right) - x_j^0 &= 0 \\ \pi_j^{-1}\left(s^0, p^0\right) - z_j^0 &= 0. \end{aligned} \tag{16}$$

We now assume the regularity conditions that enable us to relate the joint density of the structural errors $(\xi_{1t}, \ldots, \xi_{Jt}, \eta_{1t}, \ldots, \eta_{Jt})$ to the joint density of the observables (s_t, p_t).

Assumption 10. The random variables $(\xi_1, \ldots, \xi_J, \eta_1, \ldots, \eta_J)$ have a positive joint density $f_{\xi,\eta}$ on $\mathbb{R}^{2J}$.

Assumption 11. The vector function $\left(\sigma_1^{-1}, \ldots, \sigma_J^{-1}, \pi_1^{-1}, \ldots, \pi_J^{-1}\right)'$ has continuous partial derivatives and nonzero Jacobian determinant.

Finally, we add assumptions on the excluded demand and cost shifters. Assumption 12 requires full independence from the structural errors, while Assumption 13 ensures that these instruments have sufficient variation to trace out the demand and cost curves.

Assumption 12. $(x_t, z_t) \perp\!\!\!\perp (\xi_t, \eta_t)$.

Assumption 13. $\text{supp}(x_t, z_t) = \mathbb{R}^{2J}$.

With these assumptions, we can now show the identifiability of demand.

Theorem 4. *Suppose Assumptions 1a, 2, 3, and 8–13 hold. Then for all* j *(i)* ξ_{jt} *is identified for all* t*, and (ii) the function* $s_j\left(\mathcal{J}, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}}\right)$ *is identified on* $\chi^{\mathcal{J}}$.

Proof. We observe the joint density of market shares and prices, conditional on the vectors x_t and z_t. Under Assumptions 8–12 this joint density is related to that of $(\xi_{1t}, \ldots, \xi_{Jt}, \eta_{1t}, \ldots, \eta_{Jt})$ by

$$f_{s,p}\left(s_t, p_t | x_t, z_t\right) =$$
$$f_{\xi,\eta}\left(\sigma_1^{-1}\left(s_t, p_t\right) - x_{1t}, \ldots, \sigma_J^{-1}\left(s_t, p_t\right) - x_{Jt}, \pi_1^{-1}\left(s_t, p_t\right) - z_{1t}, \ldots, \pi_J^{-1}\left(s_t, p_t\right) - z_{Jt}\right) |\mathbb{J}(s_t, p_t)|$$

where $|\mathbb{J}\left(s_t, p_t\right)|$ is the absolute value of the Jacobian determinant for the vector function $\left(\sigma_1^{-1}, \ldots, \sigma_J^{-1}, \pi_1^{-1}, \ldots, \pi_J^{-1}\right)'$ evaluated at the point (s_t, p_t). Therefore, for any observed $(\hat{s}, \hat{p}, \hat{x}, \hat{z}, x, z)$ we can construct the ratio

$$\phi\left(\hat{s}, \hat{p}, \hat{x}, \hat{z}, x, z\right) \equiv \frac{f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s}, \hat{p}\right) - x_1, \ldots, \pi_J^{-1}\left(\hat{s}, \hat{p}\right) - z_J\right) |\mathbb{J}\left(\hat{s}, \hat{p}\right)|}{f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s}, \hat{p}\right) - \hat{x}_1, \ldots, \pi_J^{-1}\left(\hat{s}, \hat{p}\right) - \hat{z}_J\right) |\mathbb{J}\left(\hat{s}, \hat{p}\right)|}. \tag{17}$$

The Jacobian determinants cancel.[28] Thus, fixing $(\hat{s}, \hat{p}, \hat{x}, \hat{z})$, $\phi\left(\hat{s}, \hat{p}, \hat{x}, \hat{z}, x, z\right)$ is equal to the joint density $f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s}, \hat{p}\right) - x_1, \ldots, \pi_J^{-1}\left(\hat{s}, \hat{p}\right) - z_J\right)$ rescaled by the constant denominator in (17). Since this density must integrate (over $(x, z) \in \mathbb{R}^{2J}$) to one, the constant is uniquely

[28] This "trick" of using ratios of densities to cancel the Jacobian determinant is a critical step and was used by Matzkin (2005) (section 6) to sketch a constructive identification argument for a simultaneous equations model with the same form that we obtain after inverting the market share and pricing equations. The sketch uses the trick in a different way and requires, in addition to our location and scale normalizations, knowledge of the Jacobian determinant at one point. Completing the sketch would require showing that a particular system of nonlinear simultaneous equations has a unique solution; this appears to require further conditions on the density of unobservables. The formal results in Matzkin (2008) and Matzkin (2005) likewise rely on conditions we do not require.

determined and the value of

$$f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s},\hat{p}\right)-x_1,\ldots,\pi_J^{-1}\left(\hat{s},\hat{p}\right)-z_J\right)$$

is identified for any $(\hat{s},\hat{p},x,z)$. Since

$$\int_{\tilde{x}_j\geq x_j,\tilde{x}_{-j},\tilde{z}} f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s},\hat{p}\right)-\tilde{x}_1,\ldots,\pi_J^{-1}\left(\hat{s},\hat{p}\right)-\tilde{z}_J\right)d\tilde{x}d\tilde{z}=F_{\xi_j}\left(\sigma_j^{-1}\left(\hat{s},\hat{p}\right)-x_j\right) \tag{18}$$

the value of $F_{\xi_j}\left(\sigma_j^{-1}\left(\hat{s},\hat{p}\right)-x_j\right)$ is then known for any $(\hat{s},\hat{p},x_j)$. By the normalization (16), $F_{\xi_j}\left(\sigma_j^{-1}\left(s^0,s^0\right)-x_j^0\right) = F_{\xi_j}(0)$. For any (s_t,p_t) we can then find x^* such that $F_{\xi_j}\left(\sigma_j^{-1}\left(s_t,p_t\right)-x^*\right) = F_{\xi_j}(0)$, which reveals $\sigma_j^{-1}(s_t,p_t)=x^*$. This identifies the function $\sigma_j^{-1}(s_t,p_t)$. With equation (11) this identifies ξ_{jt} for all t. Repeating for all j, all ξ_{jt} are identified. Part (ii) then follows directly (see the proof of Theorem 2). □

This provides a constructive proof of the identification of demand. As with our analysis using general IV conditions, we can extend the identification of demand to obtain full identification of the random utility model under the additional restriction that utility is quasi-linear in the index δ_{jt}.

Theorem 5. *Suppose Assumptions 1b, 2, 3, and 8–13 hold. Then the joint distribution of* $(v_{i1t},\ldots,v_{iJt})$ *conditional on any* $\{x_{kt},p_{kt},\xi_{kt}\}_{k\in\mathcal{J}}\in\chi$ *is identified.*

Proof. Under the specification (4), the outside good has market share

$$\Pr\left(x_{1t}+\xi_{1t}+\mu_1\left(p_{1t},\omega_{i1t}\right)<0,\ldots,x_{Jt}+\xi_{Jt}+\mu_J\left(p_{Jt},\omega_{it}\right)<0\right)$$

which is

$$F_\mu\left(-x_{1t}-\xi_{1t},\ldots,-x_{Jt}-\xi_{Jt}\right). \tag{19}$$

Theorem 4 showed that each ξ_{jt} was identified. Under the full support assumption (Assumption 13), (19) determines F_μ on its full support. Since $u_{ijt}=x_{jt}+\xi_{jt}+\mu_j\left(p_{jt},\omega_{it}\right)$, this gives the result. □

6.2 Identification of Marginal Costs

We obtained identification of demand and of the full random utility model without a full model of supply. Without any additional assumption we can use the same argument to show identification of the cost shocks η_{jt}. We can then show identification of marginal cost if we are willing to assume a particular model of oligopoly competition.

Theorem 6. *Suppose that Assumptions 1a, 2, 3, and 8–13 hold. Then, for all j (i) each* η_{jt} *is identified and (ii) if each* ψ_j *is known, the function* $mc_j(q_{jt}, \zeta_{jt})$ *is identified on the support of* (q_{jt}, ζ_{jt}).

Proof. Part (i) follows by observing that the argument used in the proof of Theorem 4 can be repeated with trivial modification to recover the inverse pricing relations π_j^{-1} and the cost shocks ζ_{jt}.[29] Now observe that

$$mc_j(q_{jt}, \zeta_{jt}) = \psi_j\left(s_t, M_t, D_t\left(s_t, p_t\right), p_t\right). \tag{20}$$

Theorem 4 ensures that $D_t\left(s_t, p_t\right)$ is known. Thus all arguments of ψ_j are known. Since ψ_j is itself known, the right side of (20) is known. Since $q_{jt} = M_t s_{jt}$, by part (i) both arguments of the left side of (20) are known. Part (ii) then follows. □

Combining this result with those in section 6.1, we have provided conditions for identification of costs and demand. The overall argument is analogous to classical identification arguments for supply and demand models, which involve excluded demand shifters and cost shifters with sufficient support to trace out the supply and demand functions.

7 Discriminating Between Oligopoly Models

A remaining question is whether the correct model of competition can be distinguished from others. Bresnahan (1982) and Lau (1982) showed that this was possible for homogeneous

[29] Starting with equation (18), integrate instead over $\{\tilde{x}, \tilde{z}_{-j}, \tilde{z}_j \geq z_j\}$ and then use the normalization $\pi_j^{-1}\left(s^0, p^0\right) - z_j^0 = 0$.

goods markets within the context of deterministic conjectural variations models.[30] The following shows that the same is true in our environment, which allows for product differentiation, structural errors (our ξ_{jt} and η_{jt}), and models outside the conjectural variations framework.

Remark 1. *Consider any two markets t and t' such that $(q_{jt}, \zeta_{jt}) = (q_{jt'}, \zeta_{jt'})$. Under Assumption 8,* $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_j(s_{t'}, D_{t'}(s_{t'}, p_{t'}), p_{t'})$.

This remark provides a testable restriction that can be used to distinguish between alternative models of supply as long is the conditions for part (i) of Theorem 3 or part (i) of Theorem 6 hold for both models. To illustrate this, consider first the simple case of homogeneous goods/firms when one has a null hypothesis of perfect competition and an alternative hypothesis of monopoly. Suppose the truth is monopoly and consider Figure 2. Here the market demand curve is D_t and the true marginal revenue curve is MR_t. Thus in the monopoly model, the function ψ_j in Assumption 8 is the curve MR_t. We label this curve ψ^1_{jt}, indicating the alternative hypothesis. Under the null of perfect competition, however, it is the demand curve that is the function ψ_j. We label this ψ^0_{jt}. The observed equilibrium outcome E_t in market t maps to two possible values of marginal cost at the quantity q_t, depending on the model.

Now hold the cost shocks fixed—remember that these are identified without knowledge of the true model—and consider a change in market conditions that "rotates" the marginal revenue curve ψ^1_{jt} around the point (q_t, mc^1_t). This is illustrated in Figure 3 with the curve $\psi^1_{jt'}$. Associated with this new marginal revenue curve is a market demand curve $\psi^0_{jt'}$. Since the true model is monopoly, the new equilibrium point is $E_{t'}$. Under the alternative, the implied marginal cost at quantity q_t is again mc^1_t, consistent with the restriction Remark 1. However, under the null, the implied marginal cost is $mc^0_{t'}$, which is different from mc^0_t. So the restriction is violated and the false null is ruled out.

This describes a general "recipe" for ruling out false models using Remark 1: any exogenous changes in market conditions that rotate the true ψ_j will lead to a violation that rejects

[30] See also Bresnahan (1989).

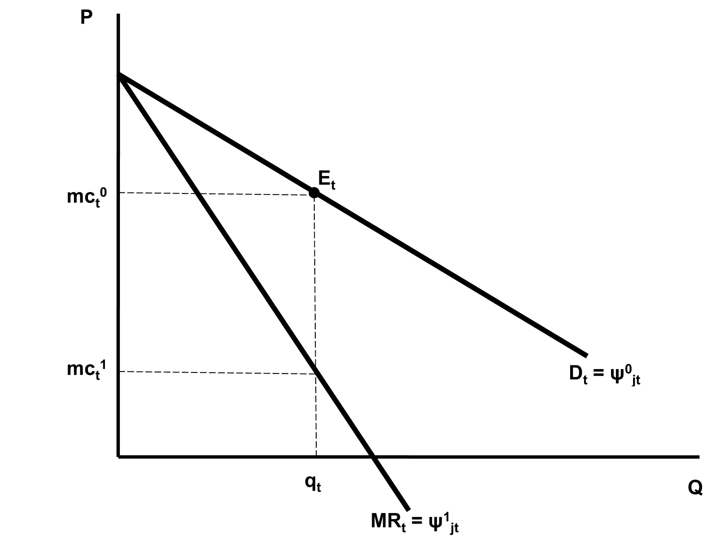

Figure 2: Market outcome E_t maps to different marginal costs under the null and alternative.

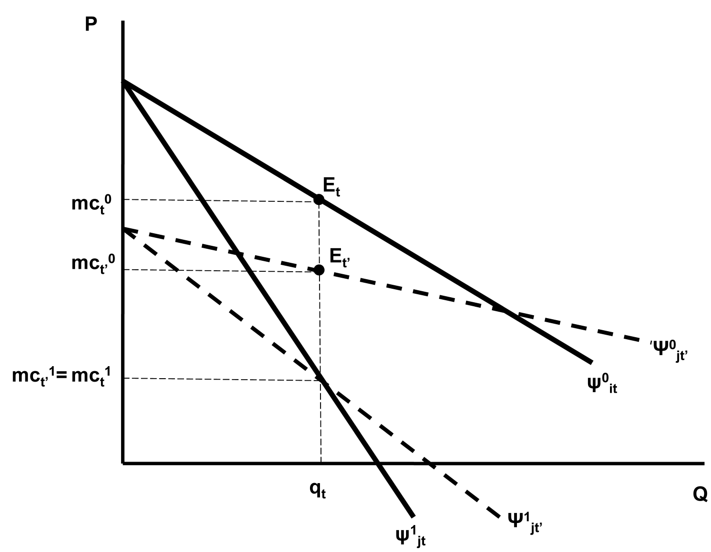

Figure 3: A rotation of the true ψ_j rules out the false null.

a false null. Indeed, the heuristic illustration in Figure 3 applies to any null and alternative oligopoly models. Any false null can be ruled out as long as there exist changes in the market environment that induce rotations under the true model that are not also rotations under the false null. This is easy to see analytically as well. Consider two markets t and t' with $\zeta_{jt} = \zeta_{jt'} = \zeta$ and $q_{jt} = q_{jt'} = q$. Since $mc_j\left(q_{jt}, \zeta_{jt}\right) = mc_j\left(q_{jt'}, \zeta_{jt'}\right) = mc_j\left(q, \zeta\right)$, we can use (20) to rationalize $q_{jt} = q_{jt'} = q$ under the null only if ψ_j is the same in both markets under the null, as it is under the true model.

This observation falls directly out of our identification analysis but is very closely related to well known insights of for the class of conjectural variations models (e.g., Bresnahan (1982), Lau (1982)). Our graphical illustration is closely related to that given by Bresnahan (1982) in particular, but makes explicit the key role of the "residual marginal revenue" function, $\psi_j(\cdot)$. The basic insight is the same: one can distinguish between competing models as long as there are changes in the market environment that can shift equilibrium quantities and markups independently, at least for some firms. Precise conditions guaranteeing such variation will depend on the model; however, the changes in the environment that alter ψ_j may form a much larger set than the changes in demand considered in the formal results of Lau (1982) or the intuitive notion of "rotations in demand" described by Bresnahan (1982, 1989). For example, even if preferences are identical in markets t and t' (i.e., there is no change in demand), ψ_{jt} and $\psi_{jt'}$ can differ due to variation in the number of competitors, the characteristics of competitors' products, or the costs of competing firms.[31]

8 Conclusion

We have considered nonparametric identification in a class of differentiated products models used in a growing body of empirical work in IO and other fields of economics. These models feature rich heterogeneity in preferences and make explicit account for endogeneity arising through market/choice-specific unobservables. Often, prices and/or other product charac-

[31] Although we have conditioned on values of $x_t^{(2)}$ and M_t, if any of these varies exogenously this variation can also be exploited.

teristics are chosen by profit-maximizing firms with unknown cost functions. Despite the wide empirical application, the sources of identification in such models has not been fully understood, leaving questions about the extent to which empirical results are fundamentally dependent on functional form and/or distributional assumptions.

We considered a nonparametric generalization of parametric random utility models used in applied work and provided positive identification results for demand, the full random utility model, and for marginal costs, using two alternative types of arguments. One argument links identification of these models to the same kinds of conditions used to show identification of regression models, while the other has close connections to parallels to identification arguments for supply and demand models. We also pointed to testable implications that can be used to discriminate between alternative models of oligopoly competition.

Our hope is that our results will be useful to both producers and consumers of empirical work on differentiated products markets. The results should help practitioners focus on the essential sources of variation. For applications in which identification of demand is sufficient, the critical issue is the availability of instruments. It should be no surprise that there is no getting around the need for instruments, but it should be comforting that this is essentially all that is needed.[32] Likewise our work should help policy makers, managers, and others who might rely on discrete choice demand estimates for making decisions. Our results demonstrate that the nonparametric foundation for empirical work based on these models is really no different from that for simpler, more familiar models. We hope this will aid critical readers in focusing on the key sources of variation in particular applications and ultimately lead to more informed decision making.

[32] An important caveat is that commonly used instruments—exogenous characteristics of competiting products (Berry, Levinsohn, and Pakes (1995)) may not be sufficient on their own without additional restrictions. A point emphasized in Berry and Haile (2009b) is that with micro data, natural instruments will often be more readily available. Combined with the stronger identification results available in the micro data environment, this provides a powerful motivation for researchers to seek individual-level data to replace or complement market level data when possible.

Appendix A

Here state an elementary result in matrix theory (see, e.g., Horn and Johnson (1990), section 6.2) that will be useful in the two appendices that follow.

Lemma 3. *Consider an $n \times n$ matrix A. The following conditions are equivalent:*
(i) the directed graph of A is strongly connected;
(ii) for any strict subset $\mathcal{K} \subset \{1, \ldots, n\}$, there exists $\ell \in \mathcal{K}$ and $\ell \notin \mathcal{K}$ such that $a_{\ell k} \neq 0$;
(iii) A is irreducible.[33]

Appendix B

Proof of Lemma 1. Suppose that for some $\delta \neq \delta'$, $\sigma_j(\delta, p) = \sigma_j(\delta', p) = s_j$ for all j. Since we have normalized the utility of the outside good to zero for all choice sets, we can define $\delta_0 = \delta_0^{'} = 0$ as a notational convention without loss. Because u_j is strictly increasing in δ_{jt}, Assumption 3 and Lemma 3 (part ii) from Appendix B imply that for any δ such that $\sigma_j(\delta, p) > 0$ for all $j \in \mathcal{J}$ and for any strict subset $\mathcal{K} \subset \mathcal{J}$, there exists $k \in \mathcal{K}$ and $\ell \notin \mathcal{K}$ such that $\sigma_\ell(\delta, p)$ is strictly decreasing in δ_k. If $\delta_j' < \delta_j \; \forall j > 0$, we must have $\sigma_0(\delta', p) > \sigma_j(\delta, p)$ (let $\mathcal{K} = \mathcal{J} - \{0\}$). Thus, we must have $\delta_m' > \delta_m$ for some $m > 0$. Because $0 \in \mathcal{J}$, there must then exist a nonempty strict subset of choices $\mathcal{K} \subset \mathcal{J}$ such that $\delta_j' > \delta_j \; \forall j \in \mathcal{K}$ and $\delta_j' \leq \delta_j \; \forall j \in \mathcal{J} - \mathcal{K}$. For this subset $\mathcal{K}$ let $k \in \mathcal{K}$ be the index of a product referred to as "k" above. Now define a new vector δ^* by

$$\begin{aligned} \delta_k^* &= \delta_k^{'} \\ \delta_j^* &= \delta_j \quad \forall j \neq k. \end{aligned}$$

[33] A matrix is *reducible* if and only if it can be placed into block upper triangular form by permutations of rows and columns. A matrix that is not reducible is *irreducible.*

Monotonicity of v_{ijt} in δ_{jt} implies that $\sigma_j(\delta^*, p) \leq \sigma_j(\delta, p)$ for all $j \in \mathcal{J} - \mathcal{K}$. Furthermore, Assumption 3 implies

$$\sum_{j \in \mathcal{J}-\mathcal{K}} \sigma_j(\delta^*, p) < \sum_{j \in \mathcal{J}-\mathcal{K}} \sigma_j(\delta, p)$$

so that, since probabilities must sum to one,

$$\sum_{j \in \mathcal{K}} \sigma_j(\delta^*, p) > \sum_{j \in \mathcal{K}} \sigma_j(\delta, p).$$

But then by monotonicity of v_{ijt} in δ_{jt}, we have

$$\sum_{j \in \mathcal{K}} \sigma_j(\delta', p) \geq \sum_{j \in \mathcal{K}} \sigma_j(\delta^*, p) > \sum_{j \in \mathcal{K}} \sigma_j(\delta, p)$$

contradicting the hypothesis $\sigma_j(\delta, p) = \sigma_j(\delta', p) = s_j$ for all j. □

Appendix C

In the text we provided a high-level condition—part (iii) of Assumption 8—ensuring that oligopoly first-order conditions can be inverted to solve for marginal cost, given the demand system. Here we show that, under assumptions already maintained in our analysis, this high-level assumption is satisfied in standard oligopoly models, including the multi-product price or quantity setting models widely used in applications. As emphasized in the text, the strategy of solving first-order conditions for marginal costs has a long history in the IO literature (e.g., Rosse (1970), Bresnahan (1981), Bresnahan (1987), and Berry, Levinsohn, and Pakes (1995)). The innovation in this appendix is the demonstration, under general nonparametric conditions, of the invertibility of particular substitution matrices. A key condition used below is the same "connected substitutes" condition we relied on to show the invertibility of the demand side of the models in Lemma 3.

We first discuss several standard models, noting the invertibility conditions that will ensure a solution for marginal cost. We then show that the structure already assumed is sufficient to ensure this invertibility in all the examples. Thus, for any of the models discussed here, part (iii) of Assumption 8 could be viewed as a lemma (proved in this appendix) rather than an assumption.

Examples of First-Order Conditions

The simplest case is the perfectly competitive model, where firms are symmetric and

$$\psi_j\left(s_t, D_t\left(s_t, p_t\right), p_t\right) = p_{jt}.$$

This provides a solution for marginal cost (marginal cost equals price) with no assumptions on demand. Of course, perfect competition is seldom a natural assumption for differentiated products markets. We therefore turn to a set of standard oligopoly models. We consider both the case of single-product firms and the more general case of multi-product firms, which also nests the case of monopoly (perfect collusion).

The most common assumption for empirical work on differentiated products markets is Nash equilibrium in a complete information simultaneous price-setting game. For single-product firms, the first-order condition is (letting $\sigma_{jt} = \sigma_j\left(s_t, p_t\right)$ as shorthand)

$$\sigma_{jt} + (p_{jt} - mc_{jt})\frac{\partial \sigma_{jt}}{\partial p_{jt}} = 0$$

which is easily solved for marginal cost:

$$mc_{jt} = p_{jt} + \frac{\sigma_{jt}}{\partial \sigma_{jt}/\partial p_{jt}}.$$

As long as the demand derivative $\partial \sigma_{jt}/\partial p_{jt}$ is non-zero, the right-hand side provides the required function $\psi_j\left(s_t, D_t\left(s_t, p_t\right), p_t\right)$.

The condition needed for the multi-product price-setting case is slightly more complicated. The first-order condition for the price of good j, produced by firm f, is

$$\sigma_{jt} + \sum_{k \in \mathcal{J}_f} (p_{kt} - mc_{kt}) \frac{\partial \sigma_{kt}}{\partial p_{jt}} = 0$$

where $\mathcal{J}_f$ is the subset of products in $\mathcal{J}$ produced by f. As in the empirical work of Bresnahan (1981) and Bresnahan (1987), the vector of first-order conditions can then be written as

$$\sigma_t + \Delta_t (p_t - mc_t) = 0 \tag{21}$$

where the (k, j) element of the square matrix Δ_t is equal to $\partial\sigma_{kt}/\partial p_{jt}$ if products k and j are produced by the same firm and equal to zero otherwise. Following BLP, the supply-side "inversion" for marginal cost is then

$$mc_t = p_t + \Delta_t^{-1} s_t.$$

In this multi-product price-setting case, to satisfy part (iii) of Assumption 8 the matrix Δ_t must be invertible.

Turning to quantity-setting models, we first require existence of an inverse demand function

$$p_t = \rho(s_t, \delta_t).$$

Consider first the case of single-product firms. Given inverse demand, the first-order condition for the simultaneous quantity setting game equates marginal cost and marginal revenue:

$$mc_{jt} = p_{jt} + \frac{\partial \rho_j}{\partial s_{jt}} s_{jt}.$$

Thus we require that the derivative $\frac{\partial \rho_j}{\partial s_{jt}}$ exist. With multi-product firms (which nests multi-product monopoly/perfect collusion), a change in the quantity of product j can change the market-clearing price for the firm's other products as well. Thus, rearranging the multi-product firm's first-order conditions gives

$$mc_{jt} = p_{jt} + \sum_{k \in \mathcal{J}_f} \frac{\partial \rho_k}{\partial s_{jt}} s_{kt}.$$

This solution requires existence of the derivatives of the inverse demand function.

Solutions for Marginal Costs

Price-Setting

For the price-setting models, we need invertibility of the within-firm substitution matrix Δ_t, which is a diagonal matrix in the case of single-product firms. We will show that invertibility is guaranteed by conditions already assumed in the text.

In the single-product price-setting case, we need existence of a nonzero derivative $\partial\sigma_{jt}/\partial p_{jt}$ for all j. This is guaranteed by Assumption 7 and part (ii) of Assumption 8.

To show invertibility of the matrix Δ_t for the multi-product case, we rely on "Taussky's theorem" which shows that an irreducibly diagonally dominant matrix is invertible.[34] A matrix A is *diagonally dominant* if

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|.$$

An *irreducibly diagonally dominant* matrix is a square matrix that is irreducible (see footnote 3) and diagonally dominant, with at least one of the diagonals being strictly dominant, i.e. with at least one row such that

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|. \tag{22}$$

Proposition 1. *For any strict subset of products $\mathcal{K} \subset \mathcal{J}$, re-index the products in $\mathcal{K}$ from 1 to $|\mathcal{K}|$ and let $D(\mathcal{K})$ be the $|\mathcal{K}|$ by $|\mathcal{K}|$ matrix with (k, j) element $\partial\sigma_{kt}/\partial p_{jt}$. Given Assumptions 3, 7 and 8, $D(\mathcal{K})$ is invertible.*

[34] See Horn and Johnson (1990), p. 363. For further background on irreducibility and dominant diagonal conditions, see chapter 6 in that text.

To show this, we begin with the following lemma:

Lemma 4. *For any strict subset $\mathcal{K} \subset \mathcal{J}$, $D(\mathcal{K})$ is a diagonally dominant matrix, with at least one strictly dominant diagonal.*

Proof. Recall that because $\sum_{k\in\mathcal{J}} \sigma_{kt} = 1$, $\sum_{k\in\mathcal{J}} \frac{\partial \sigma_{kt}}{\partial p_{jt}} = 0$. For any product $j \in \mathcal{K}$, this implies that the associated diagonal element of $D(\mathcal{K})$ satisfies

$$\left|\frac{\partial \sigma_{jt}}{\partial p_{jt}}\right| = \sum_{k\in\mathcal{K}-\{j\}} \frac{\partial \sigma_{kt}}{\partial p_{jt}} + \sum_{\ell\notin\mathcal{K}} \frac{\partial \sigma_{\ell t}}{\partial p_{jt}}. \tag{23}$$

This implies

$$\left|\frac{\partial \sigma_{jt}}{\partial p_{jt}}\right| \geq \sum_{k\in\mathcal{K}-\{j\}} \frac{\partial \sigma_{kt}}{\partial p_{jt}} \tag{24}$$

since we have required in part (ii) of Assumption 8 that v_{ij}. Furthermore, by the connected substitutes assumption, Lemma 3, and the strict monotonicity of v_{ijt} in p_{jt}, the second sum in (23) is strictly positive for at least one product $j \in \mathcal{K}$. For that j the inequality in (24) is strict. □

Proof of Proposition 1. We argue that $D(\mathcal{K})$ must be either (i) an irreducibly diagonally dominant matrix, or (ii) block-diagonal with each block being an irreducibly diagonally dominant matrix. Therefore, by Taussky's theorem, $D(\mathcal{K})$ is invertible. By Lemma 3 $D(\mathcal{K})$ is irreducible if and only if the directed graph of $D(\mathcal{K})$ is strongly connected. If the directed graph of $D(\mathcal{K})$ is strongly connected then by Lemma 4 $D(\mathcal{K})$ is an irreducibly diagonally dominant matrix and is therefore invertible. So now consider the case in which the directed graph of $D(\mathcal{K})$ is not strongly connected. By Corollary 1 and Lemma 6 (both in Appendix D) the directed graph of $D(\mathcal{K})$ can be partitioned into isolated strongly connected subgraphs. The nodes in each isolated subgraph correspond to a subset of products that do not substitute outside of the subset. We can therefore rearrange the order of products, with the products in the first strongly connected subset coming first, the next subset following and so on. The resulting permutation of $D(\mathcal{K})$ is block diagonal, with each block being irreducible by Lemma 3. Further, by Lemma 4, each block is diagonally dominant with at least one strictly dominant diagonal. Therefore, by Taussky's theorem, each block is invertible. This implies that the entire $D(\mathcal{K})$ matrix is invertible. □

We can now use Proposition 1 to prove that the matrix of within-firm derivatives, Δ_t, is invertible. First note that Δ_t is itself block diagonal with each block consisting of the $\partial\sigma_{kt}/\partial p_{jt}$ terms for the product j and k produced by a given firm. Due to the outside good, even in the case of monopoly, the set of products produced by one firm will be a strict subset of $\mathcal{J}$. Thus, by Proposition 1, each of these blocks is invertible and so the matrix Δ_t is invertible.

Quantity setting.

In the quantity setting example, the key condition is the existence of the inverse demand function, and its derivatives. We have assumed (part (ii) of Assumption 8) that v_{ijt} is strictly decreasing in p_{jt}. Thus, the same argument used to prove Lemma 1 (swapping the roles of p_t and δ_t) implies that, due to the connected substitutes property, for every (s_t, δ_t), there is a unique price vector p_t that solves $s_t = \sigma(p_t, \delta_t)$. This implies existence of an inverse demand function, which we write in vector form as $p_t = \rho(s_t, \delta_t)$.

Proposition 1 guarantees the invertibility of the matrix of own- and cross-price derivatives of market shares. So by the inverse function theorem, derivatives of the inverse demand function exist and are (as usual) equal to the elements of the inverse $D(\mathcal{K})$ matrix, i.e.,

$$\frac{\partial \rho_k}{\partial s_{jt}} = \left[D(\mathcal{K})^{-1}\right]_{kj}.$$

Appendix D

Here we present two results referenced in Appendix C. The first provides sufficient conditions for the substitution incidence matrix $\Sigma\left(J_t\right)$ to be symmetric, so that all edges of its directed graph are bidirectional. The second concerns a simple property of a strongly connected graph with bidirectional edges.

Lemma 5. *Suppose Assumptions 1a and 2 hold and that for all* $j \in J_t$ $\Pr\left(v_{ijt} \le v | x_{jt}, p_{jt}, \xi_{jt}\right)$ *is strictly decreasing and continuous in* w_{jt}.[35] *Then good* k *substitutes to good* ℓ *at* $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j \in J_t}$ *if and only if good* ℓ *substitutes to good* k *at* $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j \in J_t}$.

Proof. For good k to substitute to good ℓ, it must be the case that

$$\Pr\left(v_k - \epsilon < v_\ell < v_k\right) > 0 \quad \forall \epsilon > 0$$

i.e., (recalling that $d_{i\ell kt} \equiv v_{i\ell t} - v_{ikt}$)

$$\Pr\left\{d_{\ell k} \in (-\epsilon, 0)\right\} > 0 \;\; \forall \epsilon > 0.$$

Thus, arguing by contradiction, suppose $\Pr\left(d_{\ell k} \in (-\epsilon, 0)\right) > 0$ for some $\epsilon > 0$ but $\Pr\left(d_{\ell k} \in (0, \eta)\right) = 0$ for some $\eta > 0$. Then we must have either (a) $\Pr\{d_{\ell k} > 0\} = 0$, violating Assumption 2 , or (b) $Pr\left(d_{\ell k} > 0\right) > 0$ when supp$d_{\ell k}$ excludes $(0, \eta)$, violating Assumption 1a. □

Symmetry of $\Sigma\left(J_t\right)$ was not required by our results, but this demonstrates why in standard models (where v_{ijt} is everywhere continuos in w_{jt}) the directed graph of $\Sigma\left(J_t\right)$ is bidirectional (recall Figure 1). Further, the following Corollary 1 is utilized in Appendix C.

Corollary 1. *Consider any* $\mathcal{K} \subseteq \mathcal{J}$, *re-index the goods in* $\mathcal{K}$ *from* 1 *to* $|\mathcal{K}|$, *and let* $D(\mathcal{K})$ *be the* $|\mathcal{K}|$ *by* $|\mathcal{K}|$ *matrix with* (k, j) *element* $\partial\sigma_{kt}/\partial p_{jt}$. *Under Assumptions 1a, 2, and 8 part (ii), the directed graph of* $D(\mathcal{K})$ *is bidirectional.*

Finally, we provide the following result, referenced in Appendix C.

Lemma 6. *Suppose the directed graph* G *is strongly connected and consider any nonempty subset* N *of its nodes. If all edges of* G *are bidirectional,* N *can be partitioned into isolated strongly connected subgraphs.*

[35] As in Definition 1, w_{jt} may be an element of $\left(x_{jt}, p_{jt}, \xi_{jt}\right)$ or and index derived from $\left(x_{jt}, p_{jt}, \xi_{jt}\right)$. The relevant cases in the text are $w_{jt} = \delta_{jt} = x_{jt}^{(1)} + \xi_{jt}$ and $w_{jt} = -p_{jt}$.

Proof. Define $C(n_0) = \emptyset$, $N_0 = N$, and consider an iterative argument. Begin with iteration $t = 1$. For iteration t, let $N_t = N_{t-1} \backslash C(n_{t-1})$ and take any node $n_t \in N_t$. Let $C(n_t)$ be the set of nodes consisting of n_t and all nodes in N_t that communicate with n_t. By bidirectionality, the subgraph comprised of nodes $C(n_t)$ is strongly connected. If $N_t \backslash C(n_t) = \emptyset$, the argument is complete. Otherwise, add 1 to t and iterate. The argument will be complete in at most $|N|$ iterations. □

References

Allenby, G. M., and P. E. Rossi (1999): "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics*, 89, 57–78.

Altonji, J., and R. L. Matzkin (2005): "Cross-Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors," *Econometrica*, 73, 1053–1102.

Athey, S., and G. W. Imbens (2007): "Discrete Choice Models with Multiple Unobserved Choice Characteristics," *International Economic Review*, 48, 1159–1192.

Bajari, P., J. Fox, K. i. Kim, and S. Ryan (2009): "The Random Coefficients Logit Model is Identified," Discussion paper, University of Minnesota.

Bayer, P., F. Ferreira, and R. McMillan (2007): "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115(5), 588–638.

Benkard, L., and S. T. Berry (2006): "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on Brown (1983) and Roehrig (1988)," *Econometrica*, 74, 1429–1440.

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 23(2), 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 60(4), 889–917.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy," *American Economic Review*, 89(3), 189–211.

——— (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market," *Journal of Political Economy*, 112(1), 68–105.

Berry, S. T., and P. A. Haile (2009a): "Identification of a Nonparametrc Generalized Regression Model with Group Effects," Discussion paper, Yale University.

——— (2009b): "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," Discussion paper, Yale University.

Berry, S. T., and A. Pakes (2007): "The Pure Characteristics Demand Model," Discussion paper, Yale University.

Blundell, R. W., and J. L. Powell (2004): "Endogeneity in Semiparametric Binary Response Models," *Review of Economic Studies*, 71, 655–679.

Bresnahan, T. (1981): "Departures from Marginal Cost Pricing in the Auomobile Industry," *Journal of Econometrics*, 17, 457–482.

——— (1982): "The Oligopoly Solution Concept is Identified," *Economics Letters*, 10, 82–97.

——— (1987): "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," *Journal of Industrial Economics*, 35, 457–482.

——— (1989): "Empirical Studies of Industries with Market Power," in *The Handbook of Industrial Organization*, ed. by R. Schamlensee, and R. Willig, no. 10 in Handbooks in Economics. North-Holland.

Briesch, R. A., P. K. Chintagunta, and R. L. Matzkin (2005): "Nonparametric Discrete Choice Models with Unobserved Heterogeneity," Discussion paper, Northwestern University.

Brown, B. (1983): "The Identification Problem in Systems Nonlinear in the Variables," *Econometrica*, 51(1), 175–96.

Bundorf, K., J. Levin, and N. Mahoney (2008): "Pricing and Welfare in Health Plan Choice," Discussion paper, Stanford University.

Caplin, A., and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), 1–23.

Capps, C., D. Dranove, and M. Satterthwaite (2003): "Competition and Market Power in Option Demand Markets," *RAND Journal of Economics*, 34(5), 737–763.

Cardon, J., and I. Hendel (2001): "Asymmetric Information in Health Care and Health Insurance Markets" Evidence from the National Medical Expenditure Survery," *RAND Journal of Economics*, 32, 408–427.

Chiappori, P.-A., and I. Komunjer (2009): "On the Nonparametric Identification of Multiple Choice Models," Discussion paper, University of California San Diego.

Chintagunta, P. K., and B. E. Honoré (1996): "Investigating the Effects of Marketing Variables and Unobserved Heterogeneity in a Multinomial Probit Model," *International Journal of Marketing Research*, 13, 1–15.

Domencich, T., and D. McFadden (1975): *Urban Travel Demand: A Behavioral Analysis*. North Holland, Amsterdam.

Eizenberg, A. (2008): "Upstream Innovation and Product Variety in the United States Home PC Market," Discussion paper, Yale University.

Fan, Y. (2008): "Market Structure and Product Quality in the U.S. Daily Newspaper Market," Discussion paper, Yale University.

Fox, J., and A. Gandhi (2009): "Identifying Heterogeneity in Economic Choice Models," Discussion paper, University of Chicago.

Gandhi, A. (2008): "On the Nonparametric Foundations of Discrete Choice Demand Estimation," Discussion paper, University of Wisconsin-Madison.

Gautier, E., and Y. Kitamura (2007): "Nonparametric Estimation in Random Coefficients Binary Choice Models," Discussion paper, Yale.

Gentzkow, M., and J. Shapiro (2009): "What Drives Media Slant? Evidence from U.S. Newspapers," *Econometrica*.

Goldberg, P. (1998): "The Effects of the Corporate Average Fuel Economy Standards in the Automobile Industry," *Journal of Industrial Economics*, pp. 1–33.

Goldberg, P., and F. Verboven (2001): "The Evolution of Price Dispersion in the European Car Market," *The Review of Economics Studies*, 68(4), 811–848.

Goldberg, P. K. (1995): "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, 63(4), 891–951.

Han, A. K. (1987): "Nonparametric Analysis of a Generalized Regression Model," *Journal of Econometrics*, 35, 303–316.

Hastings, J., D. Staiger, and T. Kane (2007): "Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery," Discussion paper, Yale University.

Hausman, J., and D. Wise (1978): "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46, 403–426.

Hausman, J. A. (1996): "Valuation of New Goods under Perfect and Imperfect Competitioin," in *The Economics of New Goods*, ed. by T. F. Bresnahan, and R. J. Gordon, chap. 5, pp. 209–248. University of Chicago Press, Chicago.

Ho, K. (2007): "Insurer-Provider Networks in the Medical Care Market," Discussion paper, Columbia University.

Hoderlein, S. (2008): "Endogeneity in Semiparametric Binary Random Coefficient Models," Discussion paper, Brown University.

Hong, H., and E. Tamer (2004): "Endogenous Binary Choice Mode with Median Restrictions," *Economics Letters*, pp. 219–224.

Honoré, B. E., and A. Lewbel (2002): "Semiparametric Binary Choice Panel Data Models Without Strictly Exogenous Regressors," *emet*, 70(5), 2053–2063.

Horn, R. A., and C. R. Johnson (1990): *Matrix Analysis*. Cambrige University Press.

Hotelling, H. (1929): "Stability in Competition," *Economic Journal*, 39, 41–57.

Ichimura, H., and T. S. Thompson (1998): "Maximum Likelihood Estimation of a Binary Choice Model with Random Coefficients of Unknown Distribution," *Journal of Econometrics*, 86(2), 269–95.

Lau, L. (1982): "On Identifying the Degree of Competitiveness from Industry Price and Output Data," *Economics Letters*, 10, 93–99.

Lehman, E., and J. P. Romano (2005): *Testing Statistical Hypotheses*. Springer, New York, 3 edn.

Lewbel, A. (2000): "Semiparametric Qualitative Response Model Estimation with Unknown Heteroscedasticity or Instrumental Variables," *Journal of Econometrics*, 97, 145–177.

——— (2005): "Simple Endogenous Binary Choice and Selection Panel Model Estimators," Discussion paper, Boston College.

Lustig, J. (2008): "The Welfare Effects of Adverse Selection in Privatized Medicare," Discussion paper, Boston University.

Magnac, T., and E. Maurin (2007): "Identification and Information in Monotone Binary Models," *Journal of Econometrics*, 139, 76–104.

Manski, C. F. (1985): "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator," *Journal of Econometrics*, 27, 313–333.

——— (1988): "Identification of Binary Response Models," *Journal of the American Statitical Association*, 83(403), 729–738.

Matzkin, R. L. (1992): "Nonparametric and Distribution-Free Estimation of the Binary Choice and Threshold Crossing Models," *Econometrica*, 60(2).

——— (1993): "Nonparametric Identification and Estimation of Polychotomous Choice Models," *Journal of Econometrics*, 58.

——— (2004): "Unobservable Instruments," Discussion paper, Northwestern University.

——— (2005): "Identification in Nonparametric Simultaneous Equations," Discussion paper, Northwestern University.

——— (2007a): "Heterogeneous Choice," in *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress of the Econometric Society*, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge University Press.

——— (2007b): "Nonparametric Identification," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. Leamer, vol. 6B. Elsevier.

——— (2008): "Identification in Nonparametric Simultaneous Equations," *Econometrica*, 76, 945–978.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by P. Zarembka. Academic Press, New York.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economics Theory*, 18, 301–307.

Nair, H., P. Chintagunta, and J.-P. Dube (2004): "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants," *Quantitative Marketing and Economics*, 2, 23–58.

Nevo, A. (2000): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal ofEconomics*, 31(3), 395–421.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307–42.

Newey, W. K., and J. L. Powell (2003): "Instrumental Variable Estimation in Nonparametric Models," *Econometrica*, 71(5), 1565–1578.

Petrin, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *JPE*, 110(4), 705–729.

Petrin, A., and K. Train (2009): "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, forthcoming.

Quandt, R. E. (1966): "A Probabalistic Abstract Mode Model," in *Studies in Travel Demand, Volume II*, pp. 90–113. Mathematica, Princeton, N.J.

——— (1968): "Estimation of Modal Splits," *Transportation Research*, 2, 41–50.

Rochet, J.-C., and L. A. Stole (2002): "Nonlinear Pricing with Random Participation," *The Review of Economic Studies*, 69(1), 277–311.

Roehrig, C. S. (1988): "Conditions for Identification in Nonparametric and Parametic Models," *Econometrica*, 56(2), 433–47.

Rosse, J. N. (1970): "Estimating Cost Function Parameters without using Cost Function Data: An Illustrated Methodology," *Econometrica*, 38(2), 256–275.

Rysman, M. (2004): "Competition Between Networks: A Study of the Market for Yellow Pages," *Review of Economic Studies*, 71.

Salop, S. (1979): "Monopolistic Competition with Outside Goods," *RAND Journal of Economics*, 10, 141–156.

Severini, T. A., and G. Tripathi (2006): "Some Identification Issues in Nonparametric Models with Endogenous Regressors," *Econometric Theory*.

Villas-Boas, S. B. (2007): "Vertical Relationships Between Manufacturers and Retailers: Inference With Limited Data," *Review of Economic Studies*, 74, 625–652.