IDENTIFICATION IN DIFFERENTIATED PRODUCTS MARKETS USING MARKET LEVEL DATA

By

Steven T. Berry and Philip A. Haile

January 2010
Revised May 2012

COWLES FOUNDATION DISCUSSION PAPER NO. 1744R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

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

May 19, 2012

Abstract

We present new identification results for nonparametric models of differentiated products markets, using only market level observables. We specify a nonparametric random utility discrete choice model of demand allowing rich preference heterogeneity, product/market unobservables, and endogenous prices. Our supply model posits nonparametric cost functions, allows latent costs shocks, and nests a range of standard oligopoly models. We consider identification of demand, identification of changes in aggregate consumer welfare, identification of marginal costs, identification of firms' marginal cost functions, and discrimination between alternative models of firm conduct. We explore two complementary approaches. The first demonstrates identification under the same nonparametric instrumental variables conditions required for identification of regression models. The second treats demand and supply in a system of nonparametric simultaneous equations, leading to constructive proofs exploiting exogenous variation in demand shifters and cost shifters. We also derive testable restrictions that provide the first general formalization of Bresnahan's (1981, 1982) intuition for empirically distinguishing between alternative models of oligopoly competition. From a practical perspective, our results clarify the types of instrumental variables needed with market level data, including tradeoffs between functional form and exclusion restrictions.

*We had helpful early conversations on this topic with Rosa Matzkin and Yuichi Kitamura, as well as valuable feedback from Liran Einav, Jon Levin, Tim Bresnahan, and participants in several seminars and conferences. Financial support from the National Science Foundation is gratefully acknowledged.

1 Introduction

Models of discrete choice between differentiated products play a central role in the modern industrial organization (IO) literature and are used in a wide range of other applied fields of economics.[1] Often the discrete choice demand model is combined with an oligopoly model of supply in order to estimate markups, predict equilibrium responses to policy, or test hypotheses about firm behavior. Typically 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 predictions. However, an important question is whether parametric functional form 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 examine the nonparametric identifiability of models in the spirit of Berry, Levinsohn, and Pakes (1995) (henceforth, "BLP") and a large applied literature that has followed. We focus on the common situation in which only market level data are available, as in BLP. In such a setting one observes market shares, market characteristics, product prices and characteristics, and product/market level cost shifters, but not individual choices or firm costs. We consider identification of demand, identification of changes in aggregate consumer welfare, identification of marginal costs, identification of firms' marginal cost functions, and discrimination between alternative models of firm conduct. We also provide guidance for applied work by focusing attention on the essential role of instrumental variables, clarifying the types of instruments needed in this setting, and pointing out tradeoffs between functional form and exclusion restrictions. Our primary motivation is to develop a nonparametric foundation for a class of empirical models used widely in practice. Nonetheless, our analysis may also suggest new estimation and/or testing approaches (parametric, semiparametric, or nonparametric).

On the demand side, the models motivating our work incorporate two essential fea-

[1] Applications include studies of the sources of market power (e.g., Berry, Levinsohn, and Pakes (1995), Nevo (2001)), mergers (e.g., Nevo (2000), Capps, Dranove, and Satterthwaite (2003)), welfare gains from new goods/technologies (e.g., Petrin (2002), Eizenberg (2011)), 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 (2009)), equilibrium product quality (e.g., Fan (2011)), media bias (e.g., Gentzkow and Shapiro (2009)), asymmetric information and insurance (e.g., Cardon and Hendel (2001), Bundorf, Levin, and Mahoney (12), 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)), voting (e.g., Gordon and Hartmann (2010)), and school choice (e.g., Hastings, Kane, and Staiger (2007)).

tures. One is rich heterogeneity in preferences, which allows flexibility in demand substitution patterns.[2] The second is the presence of product/market-level unobservables. Because these unobservables are known by firms and consumers, they give rise to endogeneity of prices. Only by explicitly modeling these unobservables can one account simultaneously for endogeneity and heterogeneity in preferences for product characteristics (see section 2). Both features are essential to reliable estimation of demand elasticities in differentiates products markets. Surprisingly, this combination of features has not been treated in the prior literature on identification. Indeed, although there is a large literature on identification of discrete choice models, there are no nonparametric or semiparametric identification results even for the linear random coefficients random utility model widely used in the applied literature that motivates us.

On the supply side, the empirical literature on differentiated products employs equilibrium oligopoly models, building on early insights of Rosse (1970) and Bresnahan (1981). By combining the model of oligopoly competition with estimates of demand, one can infer marginal costs and examine a range of market counterfactuals. Following BLP, recent work typically allows for latent cost shocks and unobserved heterogeneity in cost functions, but employs a parametric specification of costs.

Our results show that the primary requirement for identification is the availability of instrumental variables. It is not surprising that instruments are needed. Less obvious is what types of exclusion restrictions suffice in this setting. Focusing on the demand side, it is intuitive that identification requires instruments generating exogenous variation in choice sets, including changes in prices. For example, BLP exploited a combination of exogenous own-product characteristics, characteristics of alternative products, and additional shifters of markups and/or costs. Following Bresnahan (1981), BLP make an intuitive argument that changes in the exogenous characteristics of competing products should help to identify substitution patterns. Ignoring the shifters of markups and costs, characteristics of competing products are sometimes referred to as "BLP instruments." However, there has been no general formal statement about the role of BLP instruments—why they would they aid identification and whether they are sufficient alone to identify demand. One difficult question has been: how could product characteristics that are not excluded from the demand system help to identify demand?

We show that these BLP instruments are in fact useful. For our results they are necessary but not sufficient: we require additional instruments such as cost shifters or proxies for costs (e.g., prices in other markets). The use of only the BLP instruments (or, e.g., only cost shifters) would therefore require either a more parsimonious model or better data. Below we discuss tradeoffs between functional form and exclusion restrictions. In Berry and Haile (2010) we consider identification of demand when one has consumer-level choice data. There we obtain results requiring fewer instruments, in some cases allowing identification when only exogenous product characteristics (the "BLP instruments") are

[2] 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) and Quandt (1968).

available.

In the following, we begin with demand, positing a nonparametric random utility discrete choice model. We require one important index restriction on how product/market-specific unobservables enter preferences, but the model is otherwise very general. Our first result shows that standard nonparametric instrumental variables conditions (e.g., Newey and Powell (2003), Andrews (2011)) suffice for identification of demand. As usual, adding a requirement of quasilinear preferences then allows identification of changes in aggregate consumer welfare.

We then move to the supply side of the model. Given identification of demand, specifying a model of oligopoly competition allows identification of marginal costs through firms' first-order conditions. Identification of firms' marginal cost *functions* then obtains with the addition of standard nonparametric instrumental variables conditions.

Next we drop the nonparametric instrumental variables conditions and consider an alternative approach, combining the demand model and supply model in a system of nonparametric simultaneous equations. This approach requires some additional structure and stronger exclusion conditions, but enables us to offer constructive proofs using more transparent variation in excluded demand shifters and cost shifters. These results therefore complement those obtained using more abstract completeness conditions.

Finally, we consider discrimination between alternative models of firm "conduct,"i.e., alternative models of oligopoly competition. The nature of oligopoly competition is itself a fundamental question of modern industrial organization, and in practice the choice of supply model can have important implications for estimates, counterfactual simulations, and policy implications. We offer the first general formalization of Bresnahan's (1981, 1982) early intuition for empirically discriminating between alternative "oligopoly solution concepts."Unlike prior formal results (Lau (1982)), ours allow product differentiation, heterogeneous firms, latent shocks to demand and costs, and oligopoly models outside the problematic "conjectural variations"framework.

Together these results provide a positive message regarding the faith we may have in a large and growing body of applied work on differentiated products markets. This message is not without qualification. In addition to the index restrictions, our results require instruments with sufficient variation on both the demand and cost side. However, while "adequate exogenous variation"is a strong requirement for any nonparametric model, the requirement here is no stronger than for regression models. Put differently, the functional form and distributional assumptions typically used in practice play their usual roles: approximation in finite samples and compensation for the gap between the exogenous variation available in practice and that required to discriminate between all nonparametric models.

To our knowledge, we provide the first and only results on the nonparametric identification of market-level differentiated products models of the sort found in BLP and related applications in IO. However, our work is connected to several theoretical and applied literatures. In the following section we briefly place our work in the context of the prior literature. We then set up the model in section 3. Our analysis based on nonparametric instrumental variable conditions is presented in section 4. Section 5 develops our alternative simultaneous equations approach. We take up discrimination between oligopoly models in section 6. Section 7 provides a discussion focused on tradeoffs between functional form and exclusion restrictions. We conclude in section 8.

2 Related Literature

There is a large body of work on identification of discrete choice models.[3] Much of that literature considers models allowing for heterogeneous preferences through a random coefficients random utility specification, but ruling out endogeneity. Ichimura and Thompson (1998) study a linear random coefficients binary choice model. Briesch, Chintagunta, and Matzkin (2005) consider multinomial choice, allowing some generalization of a linear random coefficients model. In contrast to this literature, our formulation of the underlying random utility model specifies only random utilities, not random parameters that interact with observables to generate random utilities. This allows us to substantially relax functional form and distributional assumptions relied on in earlier work.

Also essential to our demand model is the endogeneity of prices. Several papers address the identification of discrete choice models with endogeneity. Examples include Lewbel (2000), Honoré and Lewbel (2002), Hong and Tamer (2004), Blundell and Powell (2004), Lewbel (2005), and Magnac and Maurin (2007). These 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 fundamentally, add the important distinction between market/choice-specific unobservables and individual heterogeneity in preferences. This distinction allows the model to define comparative statics that account for both heteroskedasticity (heterogeneity in tastes for characteristics) and endogeneity.[4] For example, to define a demand elasticity one must quantify the changes in market shares resulting from an exogenous change in price. Accounting for heterogeneity in consumers' marginal rates of substitution between income and other characteristics requires allowing the price change to affect the covariance matrix (and other moments) of utilities. On the other hand, controlling for endogeneity requires holding fixed the market/choice-specific unobservables. Meeting both requirements is impossible in models with a single composite error for each product.

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

[4]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).

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)).[5] However, standard models of oligopoly pricing in differentiated products markets imply that each equilibrium price depends in general on the entire vector of demand shocks (and typically the vector of cost shocks as well). This rules out a triangular structure. Nonetheless, some of our results use a related strategy of inverting a multiproduct supply and demand system to recover the entire vector of shocks to costs and demand. This could be interpreted as a generalization of the control function approach. It is also related to the small literature on nonparametric simultaneous equations. Indeed, we show that our demand and supply model can be transformed into a system of simultaneous equations with a general form recently explored by Matzkin (2008) and Matzkin (2010). Applying Matzkin's results would require both a large support assumption and restrictions on the joint density of the structural errors. We show that this is unnecessary: the density restriction can be dropped, or it can be modified in order to drop the large support condition. The former result is new to the literature on nonparametric simultaneous equations while the latter was first shown in Berry and Haile (2011).

In the literature on oligopoly supply, Rosse (1970) introduced the idea of using first-order conditions for imperfectly competitive firms to infer their marginal costs from prices and demand parameters, with important further developments in, e.g., Bresnahan (1981) and BLP. We follow a similar strategy but show that parametric restrictions are not required for identification. Our insights regarding discrimination between alternative oligopoly models are related to ideas from the early empirical IO literature on inferring firm conduct from market outcomes. Bresnahan (1982), in particular, provided influential intuition for how "rotations of demand"could distinguish between alternative oligopoly models. While Bresnahan's intuition was very general, formal results (Lau (1982)) have been limited to deterministic homogeneous goods conjectural variations models (with linear "conjectures"), and have required shifters of aggregate demand. Our results avoid these restrictions and point to the more general role of rotations of a product-specific marginal revenue function.

Turning to recent unpublished papers, Berry and Haile (2009) explores related ideas in the context of a "generalized regression model"(Han (1987)). Berry and Haile (2010) considers identification of discrete choice models in the case of "micro data,"where one observes the choices of individual consumers as well as characteristics specific to each consumer and product. 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. A key insight in Berry and Haile (2010) is that within a market the market/choice-specific unobservables are held fixed.

[5] See also Chesher (2003) and Imbens and Newey (2009).

One can therefore learn key features of a random utility model by exploiting within-market variation—variation that is not confounded by variation in the market/choice-specific unobservables. That strategy is exploited throughout Berry and Haile (2010), but cannot be applied to market level data. Berry and Haile (2010) show that this additional variation can reduce the need for exclusion restrictions, can make additional instruments available, and can allow one to drop the index structure relied on throughout the present paper.

Concurrent work by Fox and Gandhi (2009) explores the identifiability of a discrete choice model in which consumer types are multinomial and conditional indirect utility functions are analytic.[6] More recently, Fox and Gandhi (2011) have explored an extension in which the dimension of the product/market level unobservables exceeds that of the choice set, showing that *ex ante* average features of demand can still be identified. A recent working paper by Chiappori and Komunjer (2009) considers identification using "micro data"(cf. Berry and Haile (2010)) in a random utility discrete choice model with an additive structure.

3 Demand Model

3.1 Random Utility Discrete Choice

Each consumer i in market t chooses a good from a set $\mathcal{J}_t = \{0, 1, \ldots, J_t\}$. We use the terms "good"and "product"interchangeably. A "market"consists of a continuum of consumers in the same choice environment. In practice markets are often defined geographically and/or temporally; however, other notions are permitted. For example, residents of the same city with different incomes, races, or family sizes might be split into distinct markets. Formally, a market t is defined by $(\mathcal{J}_t, \chi_t)$, where

$$\chi_t = (x_t, p_t, \xi_t)$$

represents the characteristics of products and/or markets. Observed exogenous characteristics of product j and/or market t are represented by $x_t = (x_{1t}, \ldots, x_{Jt})$, where each $x_{jt} \in \mathbb{R}^K$. The vector $\xi_t = \left(\xi_{1t}, \ldots, \xi_{J_t t}\right)$ represents unobservables at the level of the product and/or market. These may reflect unobserved product characteristics and/or unobserved variation in tastes across markets. Finally, $p_t = (p_{1t}, \ldots, p_{J_t t})$ represents endogenous observable characteristics, i.e., those correlated with the structural errors ξ_t. The restriction to one scalar endogenous observable for each good/market reflects standard practice but is not essential.[7] We refer to p_{jt} as "price,"reflecting the leading

[6] 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. Gandhi, Kim, and Petrin (2011) consider identification and estimation of a particular parametric variation on the standard BLP model.

[7] 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

case. Without loss, we henceforth condition on $\mathcal{J}_t = \mathcal{J} = \{0, 1, \ldots, J\}$. Let $\mathcal{X}$ denote the support of χ_t.

Consumer preferences are represented with a random utility model.[8] Consumer i in market t has conditional indirect utilities

$$v_{i0t}, v_{i1t,} \ldots, v_{iJt}$$

for the goods. For simplicity we will refer to these as "utilities." Without loss we normalize utilities relative to that of good 0 for each consumer, implying

$$v_{i0t} = 0 \quad \forall i, t.$$

Any observed characteristics of good 0 in market t are then treated as market-specific characteristics common to all products $j > 0$ in that market. In applications it is common for good 0 to represent the composite "outside good," i.e., the decision to purchase none of the goods explicitly studied, and this outside good may have no observable characteristics. Our formulation allows this as well.

Conditional on χ_t, the utilities $(v_{i1t,} \ldots, v_{iJt})$ are i.i.d. across consumers and markets, with joint distribution

$$F_v\left(v_{i1t,} \ldots, v_{iJt} | \chi_t\right).$$

We assume $\arg\max_{j \in \mathcal{J}} \; v_{ijt}$ is unique with probability one. Choice probabilities (market shares) are then given by

$$s_{jt} = \sigma_j\left(\chi_t\right) = \Pr\left(\arg\max_{k \in \mathcal{J}} \; v_{ikt} = j \,| \chi_t\right) \qquad j = 0, \ldots, J. \tag{1}$$

Of course, $s_{0t} = \sigma_0\left(\chi_t\right) = 1 - \sum_{j=1}^{J} \sigma_j\left(\chi_t\right)$. We assume that for all $\chi_t \in \mathcal{X}$, $\sigma_j\left(\chi_t\right) > 0$ for all $j \in \mathcal{J}$.[9] Let $s_t = (s_{1t}, \ldots, s_{Jt})$ and $\sigma\left(\chi_t\right) = \left(\sigma_1\left(\chi_t\right), \ldots, \sigma_J\left(\chi_t\right)\right)$.

3.2 An Index Restriction

So far the only restriction placed on the random utility model is the restriction to a scalar unobservable ξ_{jt} for each t and $j = 1, \ldots, J$. Although this is standard in the literature, it is an important restriction.[10] We now add an important index restriction. Partition

---

remain. For an empirical analysis with multiple endogenous characteristics, see Fan (2011).

[8] See, e.g., Block and Marschak (1960), Marschak (1960), McFadden (1974), and Manski (1977) for pioneering work.

[9] This amounts to treating goods with choice probability zero in market t as unavailable in that market. It should be clear that at most a bound could be obtained on the distribution of utility for a good with zero market share.

[10] A need to have no more than one structural error per observed choice probability may not be surprising.

x_{jt} as $\left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, with $x_{jt}^{(1)} \in \mathbb{R}$. Let $x_t^{(1)} = \left(x_{1t}^{(1)}, \ldots, x_{Jt}^{(1)}\right)$ and $x_t^{(2)} = \left(x_{1t}^{(2)}, \ldots, x_{Jt}^{(2)}\right)$. Define the linear indices

$$\delta_{jt} = x_{jt}^{(1)} \beta_j + \xi_{jt} \quad j = 1, \ldots, J \tag{2}$$

and let $\delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$.

Assumption 1. $F_v(\cdot|\chi_t) = F_v\left(\cdot|\delta_t, x_t^{(2)}, p_t\right)$.

Assumption 1 requires that ξ_t and $x_t^{(1)}$ affect the distribution of utilities only through the indices δ_t. With this assumption, variation in p_{jt} and $x_{jt}^{(2)}$ can have arbitrary effects on the way variation in ξ_{jt} changes the distribution of v_{ijt}, but $x_{jt}^{(1)}$ and ξ_{jt} are perfect substitutes. Put differently, the marginal rate of substitution between a unit of the characteristic measured by x_{jt} and that measured by ξ_{jt} must be constant. This linear structure is actually stronger than necessary, and we show in Appendix B that identification with a nonparametric IV approach can also be obtained with nonseparable indices $\delta_{jt} = \delta_j\left(x_{jt}^{(1)}, \xi_{jt}\right)$. The essential requirement is that $x_{jt}^{(1)}$ and ξ_{jt} enter through an index that is strictly monotonic in ξ_{jt}.

Example 1. *In applied work it is common to generate the conditional joint distributions $F_v(\cdot|\chi_t)$ from an analytical form for random utility functions*

$$v_{ijt} = v_j(x_t, \xi_t, p_t; \theta_{it})$$

where θ_{it} is a finite-dimensional parameter. The most common form is a linear random coefficients model, e.g.,

$$v_{ijt} = x_{jt}\beta_{it} - \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{3}$$

with $\theta_{it} = (\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ independent of χ_t and i.i.d. across consumers and markets. Endogeneity of price is reflected by correlation between p_{jt} and ξ_{jt} conditional on x_{jt}. The specification (3) generalizes the BLP model by dropping its parametric distributional assumptions and allowing correlation among the components of θ_{it}.[11] *In this model, a sufficient condition for Assumption 1 is that one component of x_{jt} have a degenerate coefficient, an assumption made in most applications. Compared to this example, our model relaxes several restrictions. We do not require the linearity of (3), finite-dimensional θ_{it}, θ_{it} independent of χ_t, or even the exclusion of $(x_{kt}, \xi_{kt}, p_{kt})$ from $v_j(\cdot)$ for $j \neq k$. Further, while our assumptions below will imply that demand for good j be increasing in δ_{jt} at the market level, we do not require that $x_{jt}^{(1)}$ and ξ_{jt} be "vertical"' characteristics. For example, $x_{jt}^{(1)}$ and ξ_{jt} could be observed and unobserved factors contributing to a "horizontal"characteristic (say, the acceleration capacity of a car) which consumers as a whole like, but which some consumers dislike.*

[11] In the BLP model, price enters through a nonlinear interaction with random coefficients. Since the distinction has no substantive implication for our purposes, we ignore this and refer to the model with linear interactions as the BLP model.

Because the unobservables ξ_{jt}, $j = 1, \ldots, J$ have no natural location or scale, we must normalize them in order to have a unique representation of preferences.[12] We normalize the scale by setting $\beta_j = 1 \ \forall j$, yielding

$$\delta_{jt} = x_{jt}^{(1)} + \xi_{jt}.$$

This leaves a location normalization on each ξ_{jt} to be made later.

Henceforth we condition on an arbitrary value of $x_t^{(2)}$ and suppress it in the notation. For simplicity we then let x_{jt} represent $x_{jt}^{(1)}$. Given Assumption 1, we can then rewrite (1) as

$$s_{jt} = \sigma_j\left(\delta_t, p_t\right) \qquad j = 0, \ldots, J. \tag{4}$$

3.3 Connected Substitutes

Discrete choice models inherently involve choice among weak gross substitutes.[13] For example, if v_{ijt} is strictly decreasing in p_{jt}, a fall in p_{jt} will (all else equal) raise the market share of good j and (weakly) lower market shares of all other goods. With the following "connected substitutes"assumption (Berry, Gandhi, and Haile (2012)), we will strengthen this notion of choice among substitute goods in two ways. First, we will require that the index δ_{jt} also "act like (minus) price"in the sense that a rise in δ_{jt} will (all else equal) *weakly* lower the market shares of all goods $k \neq j$. Second, we will require a minimal degree of *strict* substitution among the goods.

Assumption 2 (connected substitutes). *For $\lambda_t = -p_t$ and for $\lambda_t = \delta_t$,*
(i) $\sigma_k\left(\delta_t, p_t\right)$ is nonincreasing in λ_{jt} for all $j > 0$, $k \neq j$, and any $\left(\delta_t, p_t\right) \in \mathbb{R}^{2J}$;
(ii) for each $\left(\delta_t, p_t\right) \in$supp$(\delta_t, p_t)$ and any $\mathcal{K} \subseteq \{1, \ldots, J\}$, there exist $k \in \mathcal{K}$ and $j \notin \mathcal{K}$ such that $\sigma_k\left(\delta_t, p_t\right)$ is strictly decreasing in λ_{jt}.

Part (i) requires the goods to be weak gross substitutes in p_t and in δ_t.[14] Part (ii) requires some strict substitution as well—loosely speaking, enough to justify treating $\mathcal{J}$ as the relevant choice set. In particular, if part (ii) fails there would be some strict subset of goods $\mathcal{K}$ that substitute only to other goods in $\mathcal{K}$. A sufficient condition for Assumption 2 is that each v_{ijt} be strictly increasing in δ_{jt} and $-p_{jt}$, but unaffected by (δ_{kt}, p_{kt}) for $k \neq j$. Conditional on the index restriction, this would be standard. However, this is stronger than necessary; for example, while Assumption 2 implies that

[12] For example, let $\tilde{\xi}_t = \alpha + \xi_t \beta$ and $\tilde{F}_v\left(\cdot|x_t, p_t, \tilde{\xi}_t\right) = F_v\left(\cdot|x_t, p_t, \xi_t\right) \ \forall \left(x_t, p_t, \xi_t\right)$. Then for every $(\alpha, \beta) \in \mathbb{R}^J \times \mathbb{R}_+^J$ we have a different representation of the same preferences.

[13] Complements can be accommodated by defining the "products"as bundles of the individual goods (e.g., Gentzkow (2004)).

[14] Our use of weak gross substitutes rather than the slightly stronger notion of weak substitution in Berry, Gandhi, and Haile (2012) exploits their Proposition 1 and the assumption that $\sigma\left(\delta_t, p_t\right)$ is defined on $\mathbb{R}^{2J}$ regardless of χ. The latter seems uncontroversial here and plays no other role.

the aggregate choice probability $\sigma_j(\delta_t, p_t)$ is strictly increasing in δ_{jt} (see Berry, Gandhi, and Haile (2012)), Assumption 2 allows an increase in δ_{jt} to lower v_{ijt} for a positive measure of consumers. Thus, x_{it} and ξ_{jt} need not be vertical characteristics.

Berry, Gandhi, and Haile (2012) provide additional discussion of this assumption and demonstrate that it captures a feature common among standard models of differentiated products demand. This will be a key condition allowing us to "invert"our supply and demand models.

3.4 Observables

Let M_t denote the measure of consumers in market t and let z_t denote a vector of instruments excluded from the demand model (we discuss exclusion restrictions below). In the market data environment, the observables include $(M_t, x_t, p_t, s_t, z_t)$. When we add a model of supply we will sometimes add to this list a vector of exogenous cost shifters w_t and instruments y_t excluded from firms' marginal cost functions. Allowable overlap between elements of $(M_t, \mathrm{w}_t, x_t, y_t, z_t)$ will be characterized below.

4 Identification under Nonparametric IV Conditions

4.1 Identification of Demand

We first consider identification of the demand system, i.e., of the mapping

$$\sigma(\cdot) = (\sigma_1(\cdot), \ldots, \sigma_J(\cdot)) : \mathcal{X} \to \triangle^J$$

where $\triangle^J$ is the unit J-simplex. A key step in our identification argument is the inversion of the system of share functions in (4). The following result, shown in Berry, Gandhi, and Haile (2012), generalizes a well-known invertibility result for linear discrete choice models in Berry (1994).

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 1 and 2 there is at most one vector δ such that $\sigma_j(\delta, p) = s_j \ \forall\ j$.*

With this result we can write

$$\delta_{jt} = \sigma_j^{-1}(s_t, p_t) \quad j = 1, \ldots, J \tag{5}$$

or

$$x_{jt} + \xi_{jt} = \sigma_j^{-1}(s_t, p_t) \quad j = 1, \ldots, J. \tag{6}$$

This equation provides a reminder of the need to normalize the location of each ξ_{jt}. In this section we choose the normalization

$$E\left[\xi_{jt}\right] = 0 \quad j = 1, \ldots, J.$$

Equation (6) has a form similar to, but different from, a standard nonparametric regression model with endogenous regressors and a linear structural error. The standard model takes the form $y_t^{(1)} - \phi\left(y_t^{(2)}\right) = \epsilon_t$, where $(y_t^{(1)}, y_t^{(2)})$ are endogenous and $\phi(\cdot)$ is unknown. Newey and Powell (2003) showed identification of $\phi(\cdot)$ using excluded instrumental variables with dimension at least equal to that of $y_t^{(2)}$. In similar notation, our inverted demand model (6) takes the form $\phi_j\left(y_{jt}^{(1)}, y_{jt}^{(2)}\right) - x_{jt} = \epsilon_{jt}$ for each j. Here all endogenous variables appear inside the unknown transformation ϕ_j, and an essential instrument, x_{jt}, is not excluded. This precludes direct application of the Newey-Powell result. However, we will see that following their identification *proof* will demonstrate identification in our setting, using as instruments the exogenous $x_t \in \mathbb{R}^J$ in combination with excluded instruments of dimension at least equal to that of $y_t^{(2)}$.

Before stating the instrumental variables conditions, recall that in addition to the exogenous product characteristics x_t, the vector z_t represents instruments for p_t excluded from the determinants of $(v_{i1t}, \dots v_{iJt})$. Standard excluded instruments include cost shifters (e.g., input prices) or proxies for cost shifters such as prices of the same good in other markets (e.g., Hausman (1996), Nevo (2001)). In some applications, characteristics of consumers in "nearby" markets will be appropriate instruments (e.g., Gentzkow and Shapiro (2009), Fan (2011)). We make the following exclusion and completeness assumptions.

Assumption 3. *For all* $j = 1, \dots, J$, $E[\xi_{jt}|z_t, x_t] = 0$ *almost surely.*

Assumption 4. *For all functions* $B(s_t, p_t)$ *with finite expectation, if* $E[B(s_t, p_t)|z_t, x_t] = 0$ *almost surely then* $B(s_t, p_t) = 0$ *almost surely.*

Assumption 3 is a standard exclusion restriction, requiring mean independence between the instruments and the structural error ξ_{jt}.[15] Assumption 4 requires completeness of the joint distribution of (z_t, x_t, s_t, p_t) with respect to (s_t, p_t).[16] This is a nonparametric analog of the standard rank condition for linear models. In particular, it 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 others through the exogenous variation in the instruments. See Newey and Powell (2003), Severini and Tripathi (2006), Andrews

[15] Observe that we do not require any restriction on the joint distribution of $x_t^{(2)}$ and ξ_t. Thus our description of $\left(x_t^{(1)}, x_t^{(2)}\right)$ as "observed exogenous characteristics," while consistent with standard practice, suggests less flexibility than what is actually permitted. For example, if if ξ_{jt} and $x_{jt}^{(2)}$ are jointly determined, this presents no problem for the identification of price elasticities as long as $x_{jt}^{(1)}$ is independent of ξ_{jt} conditional on $x_{jt}^{(2)}$. We provide an example in Appendix C.

[16] Identification with weaker forms of completeness follow from the same argument used below. L^2-completeness (Andrews (2011)) would suffice under the mild restriction that the sum $x_{jt} + \xi_{jt}$ has finite variance $\forall j$. If ξ_{jt} and x_{jt} were assumed to have bounded support $\forall j$, bounded completeness would suffice.

(2011) and references therein for helpful discussion and examples. We emphasize that we require both the excluded instruments z_t and the exogenous demand shifters x_t. The "BLP instruments"x_t alone cannot satisfy Assumption 4.

Newey and Powell (2003) used analogs of Assumptions 3 and 4 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 setting.

Theorem 1. *Under Assumptions 1–4, for all $j = 1, \ldots, J$ (i) ξ_{jt} is identified with probability one for all t, and (ii) the function $\sigma_j(\chi_t)$ is identified on $\mathcal{X}$.*

Proof. For any j, rewriting (6) and taking expectations conditional on w_t, x_t, we obtain

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

so that by Assumption 3,

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

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

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

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

$$E\left[B(s_t, p_t)|z_t, x_t\right] = 0 \qquad a.s.$$

But by Assumption 4 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 with probability one by (6), proving part (i). Because choice probabilities are observed and all arguments of the demand functions $\sigma_j(\chi_t)$ are now known, part (ii) follows immediately. □

In Appendix B we show that if we strengthen the two instrumental variables assumptions as in Chernozhukov and Hansen (2005), we can generalize the model to allow

$$\delta_{jt} = \delta_j\left(x_{jt}, \xi_{jt}\right)$$

where each δ_j is an unknown function that is strictly increasing in its second argument. This avoids the linearity of the index, the requirement that x_{jt} and ξ_{jt} be perfect substitutes, and the monotonicity of $\sigma_j(\chi_t)$ in x_{jt}. Thus, the essential requirement is strict monotonicity of the index δ_{jt} in the latent demand shock ξ_{jt}.

4.2 Welfare and the Random Utility Model

The demand system $\sigma(\delta_t, s_t)$ is the only primitive from the consumer model required for most purposes motivating estimation in differentiated products markets. The important exception is evaluation of changes in consumer welfare.[17] As usual, a valid measure of aggregate consumer welfare changes can be obtained directly from demand under an additional assumption of quasilinear preferences.

Assumption 5. *$v_{ijt} = \mu_{ijt} - p_{jt}$, where the random variables $(\mu_{i1t}, \ldots, \mu_{iJt})$ are independent of p_t conditional on (x_t, ξ_t).*

We use this restriction only in this section of the paper. Note that the restriction $v_{ijt} = \mu_{ijt} - p_{jt}$ would have no content alone: even with this restriction, any joint distribution $F_v(\cdot|\chi_t)$ can be replicated by appropriate choice of the joint distribution of $(\mu_{i1t}, \ldots, \mu_{iJt})|\chi_t$, since $p_t \in \chi_t$. The conditional independence adds the requirement that that, fixing (x_t, ξ_t), prices affect the joint distribution of utilities only through the linear terms. Thus, Assumption 5 is a natural notion of quasilinearity for a random utility setting.[18] If we specified $v_{ijt} = v_j(x_t, \xi_t; \theta_{it}) - p_{jt}$, where θ_{it} is a possibly infinite-dimensional random parameter (Example 1 is a special case), Assumption 5 holds when $\theta_{it} \perp\!\!\!\perp p_t$ conditional on (x_t, ξ_t).

Given Assumption 5, identification of all changes in aggregate consumer surplus follows from standard arguments. Equivalently, welfare changes can be derived from knowledge of the conditional joint distributions of the differences in gross utilities $\left\{(\mu_{ikt} - \mu_{ijt})\right\}_{k \in \mathcal{J} \setminus j}$ on the support of the corresponding price differences. With the convention $\mu_{i0t} = p_{0t} = 0$, let $\dot{\mu}_{ijt}$ denote the differences $\left\{(\mu_{ikt} - \mu_{ijt})\right\}_{k \in \mathcal{J} \setminus j}$ and let $\dot{p}_{jt}$ denote the differences $\{(p_{kt} - p_{jt})\}_{k \in \mathcal{J} \setminus j}$. Let $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$ denote the conditional distribution of $\dot{\mu}_{ijt}$. We can then state the following result.

Theorem 2. *Suppose Assumptions 1–5 hold. Then for all (x_t, ξ_t), all $j \in \mathcal{J}$, and all $\dot{\mu} \in$supp $\dot{p}_{jt}| (x_t, \xi_t)$, $F_{\dot{\mu}_j}(\dot{\mu}|x_t, \xi_t)$ is identified.*

Proof. Taking $j = 1$,

$$\begin{aligned} \sigma_1(\delta_t, p_t) &= \Pr(-v_{i1t} \leq 0, v_{i2t} - v_{i1t} \leq 0, \ldots, v_{iJt} - v_{i1t} \leq 0 | \delta_t, p_t) \\ &= F_{\dot{\mu}_1}(-p_{1t}, p_{2t} - p_{1t}, \ldots, p_{Jt} - p_{1t} | x_t, \xi_t) \\ &= F_{\dot{\mu}_1}(\dot{p}_{1t} | x_t, \xi_t) \end{aligned}$$

where the second equality follows from Assumption 5. The demand function $\sigma_1(\delta_t, p_t)$ and unobservables ξ_t are identified (Theorem 1), and p_t is independent of $(\mu_{i1t}, \ldots, \mu_{iJt})$

[17] Certain notions of welfare changes (e.g., Pareto improvements, compensating/equivalent variation) require evaluation of changes in individual utilities. It should be clear that in general these are nonparametrically unidentified from market level data. A further obstacle is that, without additional structure, changes in the distribution of utilities (e.g., $F_v(\cdot|\chi_t)$ vs. $F_v(\cdot|\chi_{t'})$ in our model) generally do not define changes in any individual utilities. Many parametric models permit an assumption that each consumer (or each set of identical consumers) is associated with a particular realization of random parameters (see Example 1). This allows identification of possible Pareto improvements or compensating/equivalent variation, even with market-level data.

[18] Under the seemingly more general specification $v_{ijt} = \mu_{ijt} - \alpha_{it} p_{jt}$, with the random coefficient $\alpha_{it} > 0$ a.s., one can normalize the scale of each consumer's utilities without loss by setting $\alpha_{it} = 1$.

conditional on (x_t, ξ_t). Identification of $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$ on supp $\dot{p}_{1t}|(x_t, \xi_t)$ then follows. An identical argument applies for $j \neq 1$. $\square$

This result also helps to shed light on the variation required to identify the joint distribution function $F_v(\cdot|\chi_t)$, i.e., to identify the random utility model itself. Identification of $F_v(\cdot|\chi_t)$ may be of interest when considering certain estimation strategies. And given the focus of the prior literature on identification/estimation of $F_v(\cdot|\chi_t)$—for example using identification of $F_v(\cdot|\chi_t)$ to demonstrate identification of demand—a brief discussion may clarify some of the gains achieved by focusing directly on identification of demand instead.

Let $F_\mu(\cdot|x_t, \xi_t)$ denote the conditional joint distribution of $(\mu_{i1t}, \ldots, \mu_{iJt})$. It is well known that for any j the function $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$ completely determines $F_{\dot{\mu}_k}(\cdot|x_t, \xi_t)$ for all k (see Thompson (1989)). Furthermore, knowledge of $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$ for one j implies knowledge of $F_v(\cdot|\chi_t)$ under Assumption 5. For example,

$$\sigma_0(\delta_t, p_t) = F_{\dot{\mu}_0}(p_t|x_t, \xi_t) = \Pr(\mu_{i1t} \leq p_1, \ldots, \mu_{iJt} \leq p_J|\mu x_t, \xi_t) = F_\mu(p_t|x_t, \xi_t) \qquad (7)$$

and under Assumption 5 $F_\mu(\cdot|x_t, \xi_t)$ completely determines $F_v(\cdot|\chi_t)$. So if

$$\text{supp } p_t|(x_t, \xi_t) \supset \text{supp } \mu_t|(x_t, \xi_t)$$

Theorem 2 implies identification of $F_v(\cdot|\chi_t)$.

With more limited price variation, Theorem 2 delivers partial identification of each $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$, i.e., point identification on the support of $\dot{p}_{jt}|(x_t, \xi_t)$. This implies that $F_v(\cdot|\chi_t)$ will be partially identified as well. This is easily seen from (7), although in this case the demand for good zero no longer characterizes all that is learned about $F_v(\cdot|\chi_t)$. Following Thompson (1989), for all j

$$\begin{aligned} F_{\dot{\mu}_j}(\dot{p}_{jt}|x_t, \xi_t) &= \Pr\left(\dot{\mu}_{jt} \leq \dot{p}_{jt}|x_t, \xi_t\right) \\ &= \Pr\left(\Gamma_j \dot{\mu}_{0t} \leq \dot{p}_{jt}|x_t, \xi_t\right) \\ &= \int 1\{\Gamma_j m \leq \dot{p}_{jt}\}\, dF_{\dot{\mu}_0}(m|x_t, \xi_t) \\ &= \int 1\{\Gamma_j m \leq \dot{p}_{jt}\}\, dF_\mu(m|x_t, \xi_t) \qquad (8) \end{aligned}$$

where $\leq$ represents the usual component-wise partial order and Γ_j is the matrix such that $\Gamma_j \dot{\mu}_{0t} = \dot{\mu}_{jt}$. When $j = 0$, (8) is equivalent to (7), so point identification of $F_\mu(\mu|x_t, \xi_t)$ is obtained at $\mu \in$supp $p_t|x_t, \xi_t$. For $j > 0$, (8) provides additional restrictions on $F_\mu(\mu|x_t, \xi_t)$ for $\mu \notin$supp $p_t|x_t, \xi_t$. As already noted, $F_\mu(\cdot|x_t, \xi_t)$ completely determines $F_v(\cdot|\chi_t)$. Thus, what is learned about $F_v(\cdot|\chi_t)$ will combine restrictions obtained from each partially identified $F_{\dot{\mu}_j}(\cdot|x_t, \xi_t)$, i.e., from each $\sigma_j(\cdot)$.

4.3 Identification of Marginal Costs

In the applied literature, adding a supply side to the model typically proceeds by specifying a functional form for firm costs and an extensive form for the competition between firms. A set of first-order conditions characterizing equilibrium prices and quantities in terms of firms' costs and features of demand (e.g., own- and cross-price elasticities) then allow recovery of firms' equilibrium marginal costs and, under additional exclusion restrictions, their marginal cost functions. Given identification of demand, it is straightforward to show that a similar approach can be applied to demonstrate identification in a nonparametric setting.

An identification approach based on first-order conditions requires differentiability of firms' profit functions with respect to their choice variables. This will be assured if the market share functions $\sigma_j(\delta_t, p_t)$ are continuously differentiable with respect to prices, and we will assume this directly. We also slightly strengthen part of the connected substitutes assumption (Assumption 2) by requiring $\partial\sigma_k(\delta_t, p_t)/\partial p_j > 0$ whenever $\sigma_k(\delta_t, p_t)$ is strictly increasing in p_j.

Assumption 6. *(i) $\sigma_j(\delta_t, p_t)$ is continuously differentiable with respect to p_k $\forall j, k \in \mathcal{J}$; (ii) for each $(\delta_t, p_t) \in$supp(δ_t, p_t) and any $\mathcal{K} \subseteq \{1, \ldots, J\}$, there exist $k \in \mathcal{K}$ and $j \notin \mathcal{K}$ such that $\partial\sigma_k(\delta_t, p_t)/\partial p_{jt} > 0$.*

Let mc_{jt} denote the equilibrium marginal cost of production of good j in market t. For now we place no restriction on the structure of firm costs. However, we follow the literature in assuming one has committed to a model of supply, allowing each marginal cost to be expressed as a known function of equilibrium prices, equilibrium quantities (determined by s_t and M_t), and the first-order derivatives of the demand system σ.[19]

Assumption 7a. *For each $j = 1, \ldots, J$ there exists a known function ψ_j such that for any M_t and any equilibrium value of (s_t, p_t)*

$$mc_{jt} = \psi_j\left(s_t, M_t, D_t\left(s_t, p_t\right), p_t\right) \tag{9}$$

where $D_t(s_t, p_t)$ is the $J \times J$ matrix of partial derivatives $\left[\frac{\partial\sigma_k\left(\left(\sigma_1^{-1}(s_t,p_t),\ldots,\sigma_J^{-1}(s_t,p_t)\right),p_t\right)}{\partial p_\ell}\right]_{k,\ell}$.

Although this is a high-level assumption, the following remark (proved in Appendix A) demonstrates that, given differentiability and our connected substitutes assumption on demand, Assumption 7a holds in the standard models of oligopoly supply typically considered in the empirical literature, where ψ_j can be interpreted as a "product-specific marginal revenue function"for product j. Relying on this high-level assumption thus allows us to provide results for a variety of supply models at once.

[19] Our ability to express the matrix of derivatives in terms of (s_t, p_t) exploits the invertibility result of Lemma 1.

Remark 1. *Suppose supply is characterized by any of the following models, each allowing single-product or multi-product firms:*
(i) marginal cost pricing;
(ii) monopoly pricing (or joint profit maximization);
(iii) Nash equilibrium in a complete information simultaneous price-setting game;
(iv) Nash equilibrium in a complete information simultaneous quantity-setting game.
Then under Assumptions 1, 2 and 6, Assumption 7a holds.

This remark demonstrates that Assumption 7a describes a feature common to standard models of oligopoly supply. This feature immediately allows identification of marginal costs.

Theorem 3. *Suppose Assumptions 1–4, 6, and 7a hold. Then mc_{jt} is identified for all $t, j = 1, \ldots, J$.*

Proof. Immediate from (9) and Theorem 1. □

Theorem 3 shows that marginal costs (and, therefore, markups) are identified without any additional exclusion condition or any restriction on firms' cost functions. Given any of the standard models of oligopoly supply described above, only the differentiability assumption (Assumption 6) has been added to the conditions we required for identification of demand.

Marginal costs and markups are sometimes the only objects of interest on the supply side. However, unless one assumes constant marginal costs, counterfactual questions involving changes in equilibrium quantities will require identification of firms' marginal cost *functions*. Let $\mathcal{J}_j$ denote the set of goods produced by the firm producing good j. Let $q_{jt} = M_t s_{jt}$ denote the quantity of good j produced, and let Q_{jt} denote the vector of quantities of all goods $k \in \mathcal{J}_j$. We allow each marginal cost to depend on the quantities Q_{jt}, on observable cost shifters w_{jt} (which may include or consist entirely of demand shifters), and on an unobserved cost shifter $\omega_{jt} \in \mathbb{R}$. Thus, we let

$$mc_{jt} = c_j \left(Q_{jt}, \mathrm{w}_{jt}, \omega_{jt}\right) \tag{10}$$

where the unknown function c_j may differ arbitrarily across firms.

Given Theorem 3, the values of each mc_{jt} can be treated as known, so (10) takes the form of a standard nonparametric regression equation. We have the usual endogeneity problem: firm output Q_{jt} is correlated with the marginal cost shock ω_{jt}. However, typically there are many available instruments. Any demand demand shifter excluded from the cost shifters w_{jt} is a candidate instrument. These can include observable demand shifters x_{kt} for $k \in \mathcal{J}_j$. Because standard oligopoly models imply that demand shifters for *all* goods affect every market share through the consumer choice problem and equilibrium, demand shifters x_{kt} for $k \notin \mathcal{J}_j$ are also available as instruments. Other possible instruments include observables like population that vary only at the market level and may or may not directly affect market *shares*, but do affect quantities Thus, there will often be a large number of instruments for Q_{jt}.[20] Indeed, the large number of instruments may lead to testable overidentifying restrictions, something we discuss in section 6.

With adequate instruments, identification of the marginal cost function is straightforward. One way to obtain a formal result is to assume monotonicity of c_j in the unobserved cost shifter ω_{jt} and apply the identification result of Chernozhukov and Hansen (2005) for nonparametric nonseparable instrumental variables regression models. Alternatively, one could consider the separable nonparametric specification

$$mc_{jt} = \overline{c}_j \left(Q_{jt}, \mathrm{w}_{jt}, \right) + \omega_{jt} \tag{11}$$

where identification of the unknown function $\overline{c}_j$ follows by direct application of Newey and Powell's (2003) identification result for separable nonparametric regression models.[21] We give a formal statement only for the latter case. Let y_{jt} denote instruments excluded from the own-product cost shifters, as discussed above, and make the following exclusion and completeness assumptions.

Assumption 8. $E[\omega_{jt} \,|\mathrm{w}_{jt}, y_{jt}] = 0$ *almost surely for all* $j = 1, \ldots, J$.

Assumption 9. *For all* $j = 1, \ldots, J$ *and all functions* $B\left(\mathrm{w}_{jt}, Q_{jt}\right)$ *with finite expectation, if* $E\left[B\left(\mathrm{w}_{jt}, Q_{jt}\right) |\mathrm{w}_{jt}, y_{jt}\right] = 0$ *almost surely, then* $B\left(\mathrm{w}_{jt}, Q_{jt}\right) = 0$ *almost surely.*

Theorem 4. *Suppose marginal cost takes the form in* (11) *and that Assumptions 1–4, 6, 7a, 8, and 9 hold. Then for all* $j = 1, \ldots, J$, *(i) the marginal cost functions* $\overline{c}_j\left(Q_{jt}, \mathrm{w}_{jt}, \right)$ *are identified and (ii)* ω_{jt} *is identified with probability one for all* t.

Proof. Immediate from Theorem 3 and Newey and Powell (2003). □

4.4 Identifying Cost Shocks without a Supply Model

We conclude our exploration of nonparametric instrumental variables approaches to identification by providing conditions under which the latent cost shocks ω_{jt} can be identified without specifying a particular oligopoly model. This will enable us to obtain our strongest results regarding discrimination between alternative oligopoly models in section 6. In addition, some of the structure introduced here will prove useful to our exploration of identification using nonparametric simultaneous equations in section 5.

We begin by replacing the assumption of a known supply model (Assumption 7a) with the following less restrictive condition.

---

[20] In practice it is usually assumed that only q_{jt}, not Q_{jt}, enters the cost of good j. In that case only a single excluded instrument is required.

[21] Note that unlike our prior use of a Newey-Powell inspired argument to identify demand, there is nothing nonstandard about equation (11): the left-hand side variable is endogenous and all instruments are excluded from the equation.

Assumption 7b. *For all $j = 1, \ldots, J$, there exists a function ψ_j (possibly unknown) such that for any M_t and any equilibrium value of (s_t, p_t)*

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

where $D_t \left(s_t, p_t\right)$ is the $J \times J$ matrix of partial derivatives $\left[\frac{\partial \sigma_k\left(\left(\sigma_1^{-1}(s_t,p_t),\ldots,\sigma_J^{-1}(s_t,p_t)\right),p_t\right)}{\partial p_\ell}\right]_{k,\ell}$.

Assumption 7b is identical to Assumption 7a except that the functions ψ_j now need not be known. We require only that a first-order condition of this form exist. As already noted, Appendix A shows that this exists in standard oligopoly models, including those typically considered in empirical work.[22] Violations of this condition could arise in ill-defined models (e.g., models with no equilibrium) or in models with behavior characterized by corner solutions.

With this relaxation of Assumption 7a we will require additional structure on the marginal cost function. Partition w_{jt} as $\left(\mathrm{w}_{jt}^{(1)}, \mathrm{w}_{jt}^{(2)}\right)$, with $\mathrm{w}_{jt}^{(1)} \in \mathbb{R}$, and define the "cost index"

$$\kappa_{jt} \equiv \mathrm{w}_{jt}^{(1)} \gamma_j + \omega_{jt}.$$

Parallel to our index restriction on the demand model, we now require that $\mathrm{w}_{jt}^{(1)}$ and ω_{jt} affect marginal cost only through the index κ_{jt}.

Assumption 10. *For all $j = 1, \ldots, J$, $mc_{jt} = c_j(Q_{jt}, \kappa_{jt}, \mathrm{w}_{jt}^{(2)})$, where c_j is strictly increasing in κ_{jt}.*

This is an important restriction, but one that is satisfied in many standard models.[23] As with our model of demand, we focus on the case of a linear index although only monotonicity of the index in ω_{jt} is essential.

We normalize $\gamma_j = 1$ for each j without loss. Continuing to condition on (and suppress) $x_t^{(2)}$, we now also condition on a value of $\mathrm{w}_t^{(2)} = \left(\mathrm{w}_{1t}^{(2)}, \ldots, \mathrm{w}_{Jt}^{(2)}\right)$ and suppress it in the notation by letting w_{jt} now denote $\mathrm{w}_{jt}^{(1)}$. We let $\mathrm{w}_t = \left(\mathrm{w}_{1t}, \ldots, \mathrm{w}_{Jt}\right)$.

The following lemma provides a supply-side analog to our previous result on the invertibility of demand. It shows that under Assumptions 7b and 10 there is unique vector of cost indices $(\kappa_{1t}, \ldots, \kappa_{Jt})$ consistent with any vector of equilibrium prices and market shares.

Lemma 2. *Given any (M_t, s_t, p_t), there is exactly one $(\kappa_{1t}, \ldots, \kappa_{Jt}) \in \mathbb{R}^J$ consistent with Assumptions 7b and 10.*

[22] Without changing any argument, we could generalize Assumption 7b (likewise Assumption 7a) by allowing ψ_j depend arbitrarily on the demand system $\sigma(\cdot)$. We use the simpler condition since in the examples we have considered the matrix of derivatives $D_t(s_t, p_t)$ suffices.

[23] Note that $\mathrm{w}_{jt}^{(1)}$ and ω_{jt} could be any known transformations of some other observed and unobserved cost shifters.

Proof. By Assumption 7b,

$$c_j\left(Q_{jt},\kappa_{jt}\right)=\psi_j\left(s_t,M_t,D_t\left(s_t,p_t\right),p_t\right) \tag{12}$$

and by Assumption 10, the function c_j can be inverted, yielding

$$\kappa_{jt}=c_j^{-1}\left(\psi_j(s_t,M_t,D_t\left(s_t,p_t\right),p_t);Q_{jt}\right). \tag{13}$$

Since each $q_{jt}=M_t s_{jt}$, the right-hand side is an unknown function of s_t,p_t, and M_t. □

Henceforth we condition on a value of M_t and suppress it in the notation. We then rewrite (13) as

$$\mathrm{w}_{jt}+\omega_{jt}=\pi_j^{-1}\left(s_t,p_t\right) \quad j=1,\ldots,J. \tag{14}$$

Each function π_j^{-1} involves the composition of c_j^{-1} and ψ_j; although we never defined a function π_j, we use the notation π_j^{-1} as a reminder that this represents an "inversion"of supply side equilibrium conditions.

Equation (14) takes the same form as (6), and we can use it in the same way, exploiting the exogenous cost shifters w_t and demand shifters x_t as instruments.

Assumption 11. $E[\omega_{jt}|\mathrm{w}_t,x_t]=0$ *almost surely for all* $j=1,\ldots,J$.

Assumption 12. *For all functions* $B\left(s_t,p_t\right)$ *with finite expectation, if* $E\left[B\left(s_t,p_t\right)|\mathrm{w}_t,x_t\right]=0$ *almost surely then* $B\left(s_t,p_t\right)=0$ *almost surely.*

Theorem 5. *Suppose Assumptions 1–4, 6, 7b, and 10–12 hold. Then for all* $j=1,\ldots,J$, ω_{jt} *is identified with probability one for all* t.

Proof. This follows by observing that the argument used in the proof of Theorem 1 can be applied to (14) with trivial modification to recover the inverse pricing relations π_j^{-1} and the cost shocks κ_{jt}. □

An implication of Assumption 12 (given that we have held fixed both $x_t^{(2)}$ and $\mathrm{w}_t^{(2)}$) is that

$$\left\{x_t^{(1)}\cap\left\{\mathrm{w}_t^{(1)},\mathrm{w}_t^{(2)}\right\}\right\}=\emptyset=\left\{\mathrm{w}_t^{(1)}\cap\left\{x_t^{(1)},x_t^{(2)}\right\}\right\}.$$

This was not required previously, and in practice it may rule out some demand shifters for the role of $x_t^{(1)}$. However, several common types of observable product characteristics will remain good candidates. For example, in BLP (see also Petrin (2002)) a measure of a car's fuel cost, miles per dollar, enters its demand but is excluded from its marginal cost, the latter depending on the physical measure miles per gallon. This is an example of a class of instruments involving interactions between product characteristics (which may affect marginal cost) and market characteristics (which do not). The latter might include market demographics, local climate, local distribution or transportation networks, or local prices/availability of complementary goods. Other types of candidate instruments include product characteristics produced with market-specific fixed costs, such as product-specific advertising or other marketing measures (e.g., Nevo (2001), Goeree (2008)), firm locations (e.g., Capps, Dranove, and Satterthwaite (2003)), or a firm's local offerings of complementary products (e.g., Berry and Jia (2010)). In some applications product characteristics themselves may vary without affecting firm costs due to technological constraints (e.g., satellite television reception in Goolsbee and Petrin (2004)) or other exogenous market-specific factors (climate, topography, transportation network). As these examples suggest, appropriate instruments can be found in a large share of the most prominent applications of BLP-type models. Nonetheless, the required exclusion restriction does require careful attention in practice.

5 A Simultaneous Equations Approach

In this section we consider identification of the entire demand and supply model in a system of nonparametric simultaneous equations. This alternative approach requires some additional structure, but complements the nonparametric IV approach by demonstrating that one can replace the abstract completeness conditions with more transparent conditions on excluded demand shifters and cost shifters. This leads to constructive proofs with connections to classic identification arguments for models of demand and supply in perfectly competitive markets.

Throughout this section we maintain the demand index and connected substitutes assumptions (Assumptions 1 and 2). We also maintain the cost index restriction of Assumption 10 and assume equilibrium is characterized by a (possibly unknown) set of first-order conditions. Exploiting Lemmas 1 and 2, demand and supply are then characterized by the system of $2J$ equations in (6) and (14):

$$\begin{aligned} x_{jt} + \xi_{jt} &= \sigma_j^{-1}(s_t, p_t) \quad j = 1, \ldots, J \\ \mathrm{w}_{jt} + \omega_{jt} &= \pi_j^{-1}(s_t, p_t) \quad j = 1, \ldots, J. \end{aligned}$$

These equations take a form considered previously by Matzkin (2008). However, we will offer two new results—one dropping her density restriction, the other substituting an alternative density restriction in order to drop her large support requirement.

Here a different location normalization of ξ_t and ω_t will be convenient. For each j, take arbitrary (x^0, w^0) and arbitrary (s^0, p^0) in the support of $(s_t, p_t) \mid (x^0, z^0)$ and let

$$\begin{aligned} \sigma_j^{-1}\left(s^0, p^0\right) - x_j^0 &= 0 \quad j = 1, \ldots, J \\ \pi_j^{-1}\left(s^0, p^0\right) - \mathrm{w}_j^0 &= 0 \quad j = 1, \ldots, J. \end{aligned} \tag{15}$$

Although the invertibility results above (Lemmas 1 and 2) ensure that there is a unique (δ_t, κ_t) associated with any (s_t, p_t), our reliance on a change-of-variables argument will require that this map be one-to-one. The market share functions (4) already ensure that there is exactly one vector s_t associated with any (δ_t, p_t). We will assume directly that there is also only one price vector p_t consistent with any (δ_t, κ_t).

Assumption 13. *There is a unique vector of equilibrium prices associated with any* (δ_t, κ_t).

A strong sufficient condition is that, at the true marginal cost and demand functions, the equilibrium first-order conditions have a unique solution (for prices) given any (δ, κ). However, 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 admitting multiple equilibria. If there are multiple equilibria, Assumption 13 requires an equilibrium selection rule such that the same prices p_t arise whenever (δ_t, κ_t) is the same. 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 κ_{jt}). This does not seem particularly strong, but is not something we required previously.

The change-of-variables argument also require regularity conditions enabling one to relate the joint density of the latent structural errors $(\xi_{1t}, \ldots, \xi_{Jt}, \omega_{1t}, \ldots, \omega_{Jt})$ to the joint density of the observables (s_t, p_t).

Assumption 14. (ξ_t, ω_t) *have a positive joint density* $f_{\xi,\omega}$ *on* $\mathbb{R}^{2J}$.

Assumption 15. *The function* $\left(\sigma_1^{-1}(s_t, p_t), \ldots, \sigma_J^{-1}(s_t, p_t), \pi_J^{-1}(s_t, p_t), \ldots, \pi_J^{-1}(s_t, p_t)\right)'$ *has continuous partial derivatives and nonzero Jacobian determinant* $|\mathbb{J}(s_t, p_t)|$.

Finally, we will require exclusion and support conditions on the demand and cost shifters. Assumption 16 requires full independence from the structural errors, and Assumption 17 ensures that the instruments have sufficient variation to trace out the inverse demand and supply functions. Below we will see that the full support assumption is not necessary if we instead impose a shape restriction on the distribution of the unobservables.

Assumption 16. $(x_t, \mathrm{w}_t) \perp\!\!\!\perp (\xi_t, \omega_t)$.

Assumption 17. *supp*$(x_t, \mathrm{w}_t) = \mathbb{R}^{2J}$.

We first show identification of demand, which does not require us to specify the form of oligopoly competition.

Theorem 6. *Suppose Assumptions 1, 2, 6, 7b, 10, and 13–17 hold. Then for all* $j = 1, \ldots, J$, *(i)* ξ_{jt} *is identified for all* t, *and (ii) the function* $\sigma_j(\delta_t, p_t)$ *is identified on* $\mathcal{X}$.

Proof. We observe the joint density of market shares and prices, conditional on the vectors x_t and z_t. This joint density is related to that of (ξ_t, ω_t) by

$$f_{s,p}(s_t, p_t | x_t, \mathrm{w}_t) = f_{\xi,\omega}\left(\sigma_1^{-1}(s_t, p_t) - x_{1t}, \ldots, \pi_J^{-1}(s_t, p_t) - \mathrm{w}_{Jt}\right) |\mathbb{J}(s_t, p_t)| . \quad (16)$$

Since the joint density $f_{\xi,\omega}\left(\sigma_1^{-1}(s_t, p_t) - x_{1t}, \ldots, \pi_J^{-1}(s_t, p_t) - \mathrm{w}_{Jt}\right)$ must integrate to one,

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{s,p}(s_t, p_t | x_t, \mathrm{w}_t)\, dx_t d\mathrm{w}_t = |\mathbb{J}(s_t, p_t)| .$$

Combined with (16), this implies that the value of

$$f_{\xi,\omega}\left(\sigma_1^{-1}(s_t,p_t)-x_{1t},\ldots,\pi_J^{-1}(s_t,p_t)-\mathrm{w}_{Jt}\right)$$

is identified for all $\forall s_t, p_t, x_t, \mathrm{w}_t$. Then, since

$$\int_{\tilde{\mathrm{w}},\tilde{x}_{-j},\tilde{x}_j\geq x_j} f_{\xi,\omega}\left(\sigma_1^{-1}(s,p)-\tilde{x}_1,\ldots,\pi_J^{-1}(s,p)-\tilde{\mathrm{w}}_J\right)d\tilde{x}d\tilde{\mathrm{w}}=F_{\xi_j}\left(\sigma_j^{-1}(s,p)-x_j\right) \quad (17)$$

$F_{\xi_j}\left(\sigma_j^{-1}(s_t,p_t)-x_{jt}\right)$ is identified for all s_t, p_t, x_{jt}. From the normalization (15) we know $F_{\xi_j}\left(\sigma_j^{-1}(s^0,s^0)-x_j^0\right)=F_{\xi_j}(0)$. So for any (s_t,p_t) we can find x^* (unique by Assumption 14) such that $F_{\xi_j}\left(\sigma_j^{-1}(s_t,p_t)-x^*\right)=F_{\xi_j}(0)$. This reveals $\sigma_j^{-1}(s_t,p_t)=x^*$, so the function σ_j^{-1} is identified. With equation (6) this identifies ξ_{jt} for all t. Repeating for all j, all ξ_{jt} are identified. Part (ii) then follows as in the proof of Theorem 1. □

If we are willing to assume a particular model of supply—i.e., to strengthen Assumption 7b to Assumption 7a, identification of marginal costs follows directly from the identification of demand (Theorem 6) and equation (9). Alternatively, without strengthening Assumption 7b we can apply the argument used in the proof of Theorem 6 (part (i)) to show identification of the latent cost shocks ω_{jt}. This result, given as Theorem 7, will be useful for discriminating between oligopoly models (see section 6).

Theorem 7. *Suppose the hypotheses of Theorem 6 hold. Then each ω_{jt} is identified.*

Proof. The argument used in the proof of Theorem 6 can be repeated with trivial modification to recover the inverse supply relations π_j^{-1} and the cost shocks ω_{jt}.[24] □

These results demonstrate the identifiability of demand and marginal costs using an argument with close links to classical arguments for supply and demand models: excluded cost and demand shifters with sufficient support can trace out the entire demand and supply structure. The results exploit the large support Assumption 17. This type of condition is familiar from the prior literature on discrete choice models, simultaneous equations models, and even classic supply and demand models. However, a large support assumption is not essential. The following result, which follows immediately from Theorem 3 in Berry and Haile (2011), offers an alternative.

Theorem 8. *Let $\Upsilon_t=(\xi_{1t},\ldots,\xi_{Jt},\omega_{1t},\ldots,\omega_{Jt})$ and suppose Assumptions 1, 2, 6, 7b, 10, and 13–16 hold. In addition, suppose $f_{\xi,\omega}$ is twice continuously differentiable and that for almost all (s_t,p_t) there exists $(x^*(s_t,p_t),\mathrm{w}^*(s_t,p_t))$ such that the matrix*

$$\frac{\partial^2 f_{\xi,\omega}\left(\sigma_1^{-1}(s_t,p_t)-x_1^*(s_t,p_t),\ldots,\pi_J^{-1}(s_t,p_t)-\mathrm{w}_J^*(s_t,p_t)\right)}{\partial\Upsilon_t\,\partial\Upsilon_t'}$$

is nonsingular. Then , for all $j=1,\ldots,J$ (i) ξ_{jt} is identified for all t, and (ii) the function $\sigma_j(\chi_t)$ is identified on $\mathcal{X}$.

---

[24] At equation (17), integrate instead over $\{\tilde{x},\tilde{z}_{-j},\tilde{z}_j\geq z_j\}$ and then use the normalization $\pi_j^{-1}\left(s^0,p^0\right)-\mathrm{w}_j^0=0$.

This alternative result demonstrates that there can be a tradeoff between support requirements on demand/cost shifters and shape restrictions on the distribution of the structural errors. An analogous result can be obtained for identification of the cost shocks. We omit a formal statement.

6 Discriminating Between Oligopoly Models

A central question in the empirical industrial organization literature is the nature of firm "conduct,"i.e., the form of competition between firms. Firm conduct has important implications for markups, profits, efficiency, and counterfactual predictions. Bresnahan (1982), citing formal results in Lau (1982), offered an influential insight for how "rotations of demand" could be used to discriminate between alternative models of oligopoly competition based on market level data (see also Bresnahan (1989)). The class of models considered by Lau was very limited, including only deterministic conjectural variations models, with linear conjectures, homogeneous goods, and symmetric firms. However, Bresnahan's original intuition suggested much broader applicability. Here we show that this intuition can indeed be given a formal foundation in a much more general setting. Further, in contrast to the earlier focus on rotations of demand, we point out the key role of rotation of the product-specific marginal revenue functions ψ_j. This allows a much wider range of variation in the data to be exploited.

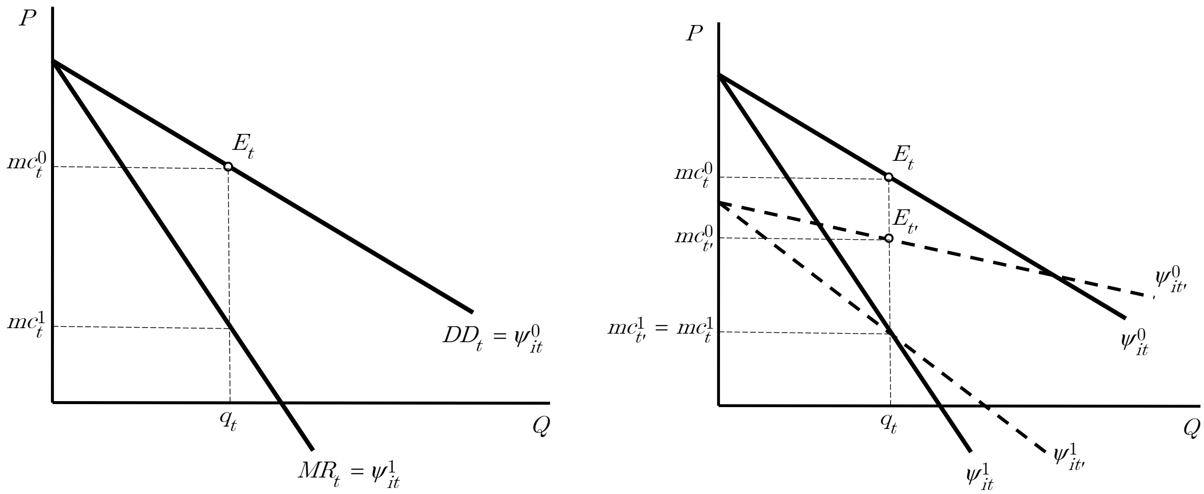
We consider two cases. We obtain our strongest results in "case 1,"where we suppose we have identified the cost indices without specifying the oligopoly model, as in Theorem 5 or Theorem 7. This case, therefore, requires both the cost index structure and exclusion of the demand shifters x_t (i.e., $x_t^{(1)}$) from firms' marginal costs. "Case 2"provides a weaker set of testable restrictions without these requirements.

6.1 Case 1

We begin with a simple example motivated by Bresnahan (1982). Consider a market with a single producer and suppose the null hypothesis is that the firm prices at marginal cost. The alternative is that the firm is a profit-maximizing monopolist. Panel (a) of Figure 1 shows the market demand curve DD_t. Under the monopoly hypothesis the function ψ_j in Assumption 7b is the marginal revenue curve MR_t. We label this curve ψ_{jt}^1, indicating the alternative hypothesis. Under the null of marginal cost pricing, however, it is the demand curve that is the function ψ_j. We label this ψ_{jt}^0. 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.

Suppose that the monopoly model is the true model. Hold the cost index fixed and consider a change in market conditions that "rotates"the curve ψ_{jt}^1 around the point (q_t, mc_t^1)—i.e., this curve changes its slope but not its level at quantity q_t. This is illustrated in panel (b) with the curve $\psi_{jt'}^1$. Associated with this new marginal revenue

Figure 1



(a) (b)

curve is a market demand curve $\psi^0_{jt'}$. Since the true model is monopoly, the new observed equilibrium outcome is $E_{t'}$. Under the alternative, the implied marginal cost at quantity q_t is still mc^1_t. Under the null, the implied marginal cost is $mc^0_{t'}$, which is different from mc^0_t. But since neither the cost index nor the quantity changed, this is impossible. This contradiction rules out the false null.

This is a particularly simple example, similar to that in Bresnahan (1982) and using a similar type of exogenous variation. But it points toward a more general strategy for discriminating between alternative models of supply based on the following result.

Theorem 9. *Suppose the function $\sigma(s_t, p_t)$ is known and that the hypotheses of either Theorem 5 or Theorem 7 hold. Then the oligopoly supply model is testable. In particular, whenever $c_j(\cdot) = c_{j'}(\cdot)$ and $(Q_{jt}, \kappa_{jt}) = (Q_{j't'}, \kappa_{j't'})$, with $t' \neq t$ and/or $j' \neq j$, the restriction $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_{j'}(s_{t'}, D_{t'}(s_{t'}, p_{t'}), p_{t'})$ must hold.*

Proof. Under Assumption 7b

$$c_j(Q_{jt}, \kappa_{jt}) = \psi_j(s_t, D_t(s_t, p_t), p_t) \ \forall j, t. \tag{18}$$

The right-hand side of (18) cannot change unless the left-hand side does. $Q_{jt}, s_t,$ and p_t are observed directly, and under the hypotheses of Theorem 5 or 7 each κ_{jt} is identified. Since $\sigma(\cdot)$ is known, $D_t(s_t, p_t)$ is known. Thus, the restriction is testable. □

In the example above, we had $(Q_{jt}, \kappa_{jt}) = (Q_{jt'}, \kappa_{jt'})$ for the single firm j and examined the restriction $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_j(s_{t'}, D_{t'}(s_{t'}, p_{t'}), p_{t'})$. With more than one good, if we assume $c_{j'}(\cdot) = c_j(\cdot)$ for some $j' \neq j$ (a common assumption is that $c_j(\cdot) = c(\cdot)$ $\forall j > 0$), then Theorem 9 requires $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_{j'}(s_t, D_t(s_t, p_t), p_t)$ whenever

$(Q_{j't}, \kappa_{j't}) = (Q_{jt}, \kappa_{jt})$. Thus, unlike the example above, the restriction can be applicable within a single market and does not require any changes in market demand. Further, there is no *a priori* restriction on the class of models that can be considered. Any false null can be ruled out as long as there exists a change across firms or across markets that induces a rotation of at least one function ψ_j under the true model that is not also a rotation under the false null.

Conditions guaranteeing such variation will depend on the models considered. But the types of variation that can alter ψ_j form a much broader class than the movements of market demand considered by Lau (1982) and Bresnahan (1982, 1989). For example, even if preferences and market size 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 competing firms, the set of competing goods, characteristics of competing products ($x^{(1)}_{-jt}$ and/or ξ_{-jt}), or costs of competing firms ($\mathrm{w}^{(1)}_{-jt}$ and/or κ_{-jt}). Since we have shown how to identify the demand shocks ξ_{jt} without knowing the oligopoly model, the variation exploited here need not be exogenous. Further, although we have conditioned on $x^{(2)}_t$, $\mathrm{w}^{(2)}_{-jt}$ and M_t, variation in any of these will typically alter ψ_j and/or D_t, and may also be exploited.

The following example illustrates some of this broader applicability in another simple but less trivial model—one closer to those used to study differentiated products markets in practice.

6.1.1 Logit Example

Consider a cross section of markets, each with two inside goods and an outside good. Demand is known to the researcher and characterized by a multinomial logit model, where $\sigma_j(\delta, p) = \frac{e^{\delta_j - p_j}}{1+e^{\delta_1 - p_1}+e^{\delta_2 - p_2}}$. For simplicity, let each firm have constant marginal costs. This is not essential and we do not assume that the researcher knows that marginal costs are constant. Supply is characterized by Nash equilibrium in a complete information simultaneous move duopoly price-setting game. Equilibrium prices are then characterized implicitly by the first-order conditions[25]

$$p_j = mc_j + \frac{1}{1-\sigma_j(\delta, p)} \quad j = 1, 2. \tag{19}$$

Let (s, p) denote one observed equilibrium outcome. Now consider another outcome (p', s') realized when all shifters of mc_1 are held fixed, mc_2 falls to mc_2' (due to an exogenous cost shifter), and the market size exogenously grows from M to M'. Because prices are strategic complements, both equilibrium prices fall. So by (19), $s_1' < s_1$. However, let the increase in market size leave

$$q_1 = Ms_1 = M's_1' = q_1'.$$

[25] Recall that in the logit model, $\frac{\partial\sigma_j(\delta,p)}{\partial p_j} = -\sigma_j(\delta, p)(1-\sigma_j(\delta, p))$ while $\frac{\partial\sigma_j(\delta,p)}{\partial p_k} = \sigma_j(\delta, p)\sigma_k(\delta, p)$.

Since both prices decline, the market share of the outside good falls, i.e., $s_0' < s_0$. This further implies $s_2' > s_2$. We will show that the two data points (p_1, p_2, s_1, s_2, M) and $(p_1', p_2', s_1', s_2', q_1', M')$ are sufficient to rule out the the most commonly considered alternative models of supply.

Marginal Cost Pricing Let $\widetilde{mc}_1$ denote the value of firm 1's marginal cost at quantity q_1 inferred under the (incorrect) hypothesis of marginal cost pricing. Since $q_1 = q_1'$, Theorem 9 requires that $\widetilde{mc}_1$ rationalize both data points, i.e.,

$$p_1 = \widetilde{mc}_1 = p_1'.$$

Since $p_1' < p_1$, this is a contradiction.

Quantity Competition With logit demand, inverse demand for good j is given by

$$p_j = \sigma_j^{-1}(s;\delta) = \delta_j + \ln(s_0) - \ln(s_j). \tag{20}$$

Since $s_0 = 1 - s_1 - s_2$, we have $\frac{\partial p_1}{\partial s_1} = -\frac{1}{s_0} - \frac{1}{s_1}$ and $\frac{\partial p_1}{\partial s_2} = -\frac{1}{s_0}$. If supply is characterized by Nash equilibrium in a complete information simultaneous move quantity-setting game, firm 1 chooses s_1 to maximize $Ms_1p_1 - C_1(Ms_1)$, where $C_1(\cdot)$ is the total cost function of firm 1. This yields the first-order condition (after substituting for $\frac{\partial p_1}{\partial s_1}$)

$$p_1 - mc_1 - \frac{s_1}{s_0} - 1 = 0. \tag{21}$$

As above, because $q_1 = q_1'$, if (21) is the correct first-order condition, one marginal cost must rationalize both data points, i.e., there must be some $\widetilde{mc}_1$ such that

$$\widetilde{mc}_1 = p_1 - \frac{s_1}{s_0} - 1 = p_1' - \frac{s_1'}{s_0'} - 1. \tag{22}$$

With logit demand $\frac{s_1}{s_0} = e^{\delta_1 - p_1}$, so (22) requires

$$p_1 - p_1' = e^{\delta_1 - p_1} - e^{\delta_1 - p_1'}$$

which is impossible because $p_1 > p_1'$ implies $e^{\delta_1 - p_1} < e^{\delta_1 - p_1'}$.

Joint Profit Maximization Finally, consider the alternative model in which prices (or quantities) are set to maximize the joint profits from the two inside goods,

$$p_1 M s_1 - C_1(Ms_1) + p_2 M s_2 - C_2(Ms_2).$$

Prices would be determined by the first-order conditions

$$\begin{aligned} 1 - (1 - s_1)(p_1 - mc_1) + s_2(p_2 - mc_2) &= 0 \\ 1 - (1 - s_2)(p_2 - mc_2) + s_1(p_1 - mc_1) &= 0. \end{aligned}$$

Straightforward algebra yields

$$\frac{1}{s_0} - (p_1 - mc_1) = 0.$$

Following the logic above, if this is the correct model we must have

$$\widetilde{mc}_1 = p_1 - \frac{1}{s_0} = p_1' - \frac{1}{s_0'}.$$

But this cannot hold, since $p_1' < p_1$ and $s_0' < s_0$.

6.2 Case 2

In Case 1, we used the fact that if ψ_j is correctly specified there must be a single cost function c_j such that equation (18) always holds—i.e., such that whenever (Q_{jt}, κ_{jt}) are the same $\psi_j(s_t, D_t(s_t, p_t), p_t)$ is the same. We now consider a version of the same idea that is applicable without the additional index and exclusion restrictions that enable identification of each κ_{jt} without specifying ψ_j. Then the restriction (18) no longer has bite when comparing pairs of observations: variation in the right-hand side can always be attributed to latent variation in κ_{jt}. However, one can obtain a testable restriction by taking expectations conditional on exogenous variables.

To demonstrate this, we will again work with the separable specification of the marginal cost function (11)

$$mc_{jt} = \bar{c}_j(Q_{jt}, \mathrm{w}_{jt},) + \omega_{jt}$$

and maintain the exclusion restrictions of Assumption 8.[26] Consider a candidate specification of ψ_j If this gives the correct first-order condition for the firm producing good j then, taking expectations of (18) using Assumption 8, we have

$$E\left[\bar{c}_j(Q_{jt}, \mathrm{w}_{jt},) \,|\mathrm{w}_{jt}, y_{jt}\right] = E\left[\psi_j(s_t, M_t, D_t(s_t, p_t), p_t) \;\, |\mathrm{w}_{jt}, y_{jt}\right] \quad a.s. \tag{23}$$

Thus, a testable restriction of the model is that there exist some (possibly counterfactual) marginal cost function $\widetilde{\bar{c}}_j(Q_{jt}, \mathrm{w}_{jt},)$ such that

$$E\left[\widetilde{\bar{c}}_j(Q_{jt}, \mathrm{w}_{jt},) \,|\mathrm{w}_{jt}, y_{jt}\right] = E\left[\psi_j(s_t, M_t, D_t(s_t, p_t), p_t) \;\, |\mathrm{w}_{jt}, y_{jt}\right] \quad a.s. \tag{24}$$

Theorem 10. *Suppose the function $\sigma(s_t, p_t)$ is known and that Assumption 8 is satisfied. Then the oligopoly supply model is testable. In particular, for each $j > 0$ there must exist a function $\widetilde{\bar{c}}_j$ consistent with (24).*

[26] A similar restriction could be derived with a nonseparable cost function and full independence of the instruments.

This can be seen as a weaker version of the restriction (18) considered in Case 1—weaker because the restriction applies only to conditional expectations rather than to each value of $\psi_j(s_t, M_t, D_t(s_t, p_t), p_t)$.[27]

Note that we did not assume the completeness condition (Assumption 9) used in Theorem 4. This is because it is possible that variation in the instruments w_{jt}, y_{jt} can rule out a false supply model even when there is no strict subset of these instruments that would point identify the marginal cost function when this is the true model. Of course, a natural special case is that in which the instruments w_{jt}, y_{jt} lead to overidentification. For example, suppose there are two distinct subsets of the instruments that would allow point identification of a marginal cost function $\widetilde{c}_j(Q_{jt}, \mathrm{w}_{jt},)$ consistent with (24) using Theorem 4 and a candidate function ψ_j. Let ω^*_{jt} denote the cost shocks implied by the first set of identifying instruments, and let ω^{**}_{jt} denote those implied by the second. A testable restriction of the candidate oligopoly specification is that $\omega^*_{jt} = \omega^{**}_{jt}$ for all j, t.

7 Discussion: Functional Form vs. Instruments

Our results demonstrate that strong functional form restrictions are not necessary for identification or for discrimination between alternative models of oligopoly competition. Rather, the essential requirement is the availability of instruments. A need for instruments to identify demand and supply is not surprising, and it should be comforting that this essentially all that is required. From a practical perspective, our results clarify the types of exclusion restrictions needed. Intuition from elementary supply and demand models might have suggested that identification of demand would require only cost shifters or other excluded instruments for prices. However, we required instruments for both prices and "quantities"(market shares), even for identification of demand alone. This need arises from the fact that in a multi-good setting, demand is defined by a system of equations, each depending on the J-dimensional endogenous price vector p_t as well as the J-dimensional latent demand shocks ξ_t. One can obtain a system of equations with one unobservable per equation when the demand system is inverted (Lemma 1) to give

$$\xi_{jt} = \sigma_j^{-1}(s_t, p_t) - x_{jt} \quad j = 1, \ldots, J. \tag{25}$$

But now each equation depends on the $2J$ endogenous variables (s_t, p_t). As we have shown, because x_{jt} is exogenous and enters with a known coefficient, it is available as one such instrument. Thus, the demand shifters x_t provide J of the $2J$ instruments required for identification, leaving a need for J additional instruments such as product-specific cost shifters.

We have seen that once demand is identified, identification of marginal costs requires no additional IV conditions, and firms' marginal cost *functions* are likely to be overidentified. Focusing then on the IV requirements for identification of demand, it is useful to

[27] The power of the restriction may be increased if one imposes shape restrictions on the marginal cost function, for example, weak monotonicity in output.

compare the inverse demand equations (25) to those in more familiar parametric models. In a multinomial logit model with linear random utilities, inverse demand takes the form (see, e.g., Berry (1994))

$$\xi_{jt} = \ln\left(\frac{s_{jt}}{s_{0t}}\right) + \alpha p_{jt} - x_{jt}\beta \tag{26}$$

or, equivalently,

$$\tilde{\xi}_{jt} = \frac{1}{\beta}\ln\left(\frac{s_{jt}}{s_{0t}}\right) + \frac{\alpha}{\beta}p_{jt} - x_{jt} \tag{27}$$

where $\tilde{\xi}_{jt} = \frac{1}{\beta}\xi_{jt}$, giving a different scale normalization of the latent demand shocks. It is clear from either equation that identification of the unknown parameters (α, β) requires, in addition to the exogenous x_{jt}, an excluded instrument for price. Valid instruments could include cost shifters or the BLP instruments x_{-jt}. Observe that in (27), as in (25), x_{jt} enters with a known coefficient; however, there it provides exogenous variation in the function of market shares, $\ln(\frac{s_{jt}}{s_{0t}})$, which enters with unknown coefficient $\frac{1}{\beta}$.

In the nested logit model, inverse demand takes the form (again see Berry (1994))

$$\xi_{jt} = \left[\ln(s_{jt}) - \ln(s_{0t}) - \lambda_g \ln(s_{j/g,t})\right] + \alpha p_{jt} - x_{jt}\beta$$

or

$$\tilde{\xi}_{jt} = \frac{1}{\beta}\left[\ln(s_{jt}) - \ln(s_{0t}) - \lambda_g \ln(s_{j/g,t})\right] + \frac{\alpha}{\beta}p_{jt} - x_{jt.} \tag{28}$$

Here λ_g is a parameter governing substitution within/between subgroups of products ("nests") and $s_{j/g,t}$ is the within-group share of product j. Relative to the more restrictive multinomial logit, one needs an additional instrument due to the presence of the endogenous variable $\ln(s_{j/g,t})$, which enters with unknown coefficient λ_g. The exogenous demand shifters x_{-jt} are available for this role, in which case x_t may again be viewed as providing the needed exogenous variation in the market shares appearing in (28).

Finally, in the BLP model, inverse demand takes the richer parametric form

$$\xi_{jt} = \tilde{\delta}_j\left(s_t, x_t, p_t, \Sigma\right) - \alpha p_{jt} - x_{jt}\beta \tag{29}$$

(or, equivalently, $\tilde{\xi}_{jt} = \frac{1}{\beta}\tilde{\delta}_j\left(s_t, x_t, p_t, \Sigma\right) - \frac{\alpha}{\beta}p_{jt} - x_{jt}$), where the function $\tilde{\delta}_j$ maps market shares to "mean utilities"$x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$. This function is known up to the parameters Σ, which govern the distribution of the random coefficients. Since (s_t, p_t) are endogenous, identification requires not only an excluded instrument for p_{jt}, but also a sufficient number of instruments (BLP instruments and/or others) to identify all parameters (α, β, Σ). This need for additional instruments arises from the presence of unknown parameters interacting with the endogenous variables (s_t, p_t).

We can see that these examples are special cases of a general parametric model

$$\xi_{jt} = \sigma_j^{-1}\left(s_t, p_t; \theta\right) - x_{jt} \tag{30}$$

where θ is a finite-dimensional parameter. Further, as additional flexibility in substitution patterns is allowed, the inverse demand function $\sigma_j^{-1}\left(s_t, p_t; \theta\right)$ takes a form that depends in a richer way on the vectors of market shares and prices, with the need for exclusion restrictions growing accordingly. Our equation (25) is the natural nonparametric extension of these examples. And as a result, our instrumental variables requirements are an extension of those in these familiar parametric models.

One message from this discussion is that there are tradeoffs between *a priori* functional form restrictions and the need for exclusion restrictions. And it suggests a greater range of alternatives that might be explored in future work by focusing directly on the form of the inverse demand system. For example, suppose (as in many empirical applications) price enters only through the demand index, so that $\delta_{jt} = x_{jt} - \alpha p_{jt} + \xi_{jt}$, with α an unknown parameter. Equation (25) then takes the semiparametric form

$$\xi_{jt} = \sigma_j^{-1}(s_t) + \alpha p_{jt} - x_{jt}.$$

Identification of demand would require only a single instrument (e.g., a market level cost shifter) beyond the exogenous product characteristics x_t. In fact, the same reduction in the need for excluded instruments can be obtained without linearity of the index. Suppose instead that

$$\delta_{jt} = \delta_j\left(x_{jt}, p_{jt}, \xi_{jt}\right)$$

where δ_j is an unknown function required only to be strictly increasing in ξ_{jt}. This is a trivial variation on the model considered in Appendix B, and applying the argument given there yields

$$s_{jt} = h_j\left(s_{-jt}, x_{jt}, p_{jt}, \xi_{jt}\right)$$

where h_j is an unknown function required only to be strictly increasing in ξ_{jt}. This equation takes the form of a nonseparable nonparametric regression model (e.g., Chernozhukov and Hansen (2005)), and identification of h_j (and therefore of each ξ_{jt}, yielding identification of demand) requires only a single excluded instrument beyond the exogenous product characteristics x_t. [28]

Finally, it is instructive to consider the case in which prices are exogenous and can therefore be fixed (as we did with $x_t^{(2)}$) and dropped from the notation to consider identification. One might speculate that identification holds without any exclusion restrictions. However, this is incorrect, again due to the determination of each quantity as a function of the entire vector ξ_t of latent demand shocks. In this case, (25) takes the form

$$\xi_{jt} = \sigma_j^{-1}(s_t) - x_{jt}.$$

The presence of the J-dimensional vector of endogenous variables s_t in the unknown function σ_j requires at least J instruments. As before, the entire J-vector x_t is available. However, no alternative instruments are available; in particular, exogenous shifters of costs (or markups) do not help since these enter only through their effect on the (fixed)

[28] Another possible restriction is the symmetry/exchangeability assumption $\sigma_j(\delta_t, p_t) = \sigma(\delta_{jt}, p_{jt}, \delta_{-jt}, p_{-jt})$ for all j. This would allow within-market (cross-product) variation to contribute to identification and would also allow a weaker version of the usual completeness condition.

prices. Thus, identification is possible here only because the index restriction leaves x_t available as instruments. This suggests not only the necessity of the BLP instruments, but the crucial role of our index restriction for making them available.

This example clarifies several other things as well. One is the fundamental difference between models like ours and those often considered in the econometrics literature on discrete choice, which lack the structural errors ξ_t. In those models, an absence of endogeneity would make identification of demand trivial in a market data setting: observed conditional choice probabilities would *be* demand. Another is the the essential role of the assumptions we do make on the random utility model. The connected substitutes assumption allows us to invert the demand system, yielding equations taking a form nearly identical to that of nonparametric regression models.[29] Monotonicity of the index δ_{jt} in the structural error ξ_{jt} ensures that each "regression"equation is monotonic in ξ_{jt} as required for existing identification results for regression models. And, as already noted, the fact that x_{jt} enters only through the index δ_{jt} yields the essential exclusion of x_{-jt} from the inverse demand function σ_j^{-1}.

8 Conclusion

We have considered nonparametric identification in a class of differentiated products models used in a large and growing body of empirical work in IO and other fields of economics. Our results demonstrate that the nonparametric foundation for these models is strong, the primary requirement being the availability of instruments providing adequate exogenous variation in the endogenous prices and quantities. Thus, functional form and/or distributional assumptions relied on in practice can be viewed as playing only their usual roles: smoothing, extrapolation, and compensation for the gap between the exogenous variation necessary to discriminate between all functions and that available in practice. Our results also provide guidance on the key sources of variation one should look for in applications, and point to tradeoffs between *a priori* functional form assumptions and the need for exclusion restrictions.

Although we motivated our analysis of demand with a widely used class of discrete choice models, most of our results hold in other types of demand models. For example, our results on the identification of demand extend immediately to any environment in which one can invert the demand system to obtain the inverse demand system (5). Berry, Gandhi, and Haile (2012) provide a range of examples. Our results on identification and falsifiability of the supply model hold also extend immediately whenever identification of demand holds.

Finally, while we have focused exclusively on identification, our analysis may suggest new estimation approaches. A critical aspect of our identification strategy is to work directly with the economic primitives of interest on the demand side of the model, in

[29]Berry, Gandhi, and Haile (2012) show that connected substitutes is "nearly necessary"for invertibility even under additional smoothness conditions not imposed here.

particular the demand functions σ_j. This approach contrasts with the usual strategy of building up a random utility model from random parameters, using this to derive demand. This approach enabled us to obtain positive results while also substantially relaxing functional form and distributional assumptions relied on previously.[30] Our alternative approach may prove useful for developing new estimators as well, whether parametric, semiparametric, or nonparametric.[31]

[30] To see one challenge in the traditional approach, consider a linear random coefficients model with no endogeneity, where $v_{ijt} = x_{jt}\beta_{it} + \epsilon_{ijt}$. Despite the linear formulation and the absence of structural errors (making identification of *demand* trivial), it is clear that the model is not identified. This is because the joint distribution of $(\beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt}|x_t)$ has dimension (conditional on x_t) greater than J, whereas behavior is determined entirely by the J-dimensional distribution $F_v(\cdot|x_t)$. Even with the assumption $\epsilon_{ijt} \perp\!\!\!\perp \beta_{it}$, the semiparametric model remains overparameterized.

[31] A recent working paper, Gandhi and Nevo (2011), pursues a semiparametric approach. Souza-Rodrigues (2011) considers a nonparametric approach for the closely related "generalized regression model."

Appendices

A Invertibility of Oligopoly First-Order Conditions

In the text we provided a high-level condition—Assumption 7b—ensuring that oligopoly first-order conditions can be inverted to solve for each marginal cost as a function of prices, market shares, and features of the demand system. Here we show that, as claimed in Remark 1, this high-level assumption is satisfied in standard models of differentiated products oligopoly under assumptions already maintained. These models include complete information games with simultaneous choice of prices or quantities, with single- or multi-product firms. Joint profit maximization (perfect collusion) is a special case. For any of the models discussed here, Assumption 7b could be viewed as a lemma (proved in this appendix) rather than an assumption.

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 is the same "connected substitutes" condition we relied on to show the invertibility of the market share function on the demand side of the model.

Throughout this appendix will fix the market size M_t and suppress it. Recall that $D_t(s_t, p_t)$ is the $J \times J$ matrix with (k, ℓ) element equal to $\frac{\partial \sigma_k\left(\sigma^{-1}(s_t,p_t),p_t\right)}{\partial p_\ell}$. Below we will make use of the follow result, which follows immediately from Theorem 2 in Berry, Gandhi, and Haile (2012).

Lemma 3. *Let Assumptions 2 and 6 hold. Then for all $\chi_t \in \mathcal{X}$, every principle submatrix of $D_t(s_t, p_t)$ is nonsingular.*

A.1 Marginal Cost Pricing

Here the result is immediate:

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

A.2 Price Setting

The most common model of supply for empirical work on differentiated products markets is Nash equilibrium in a complete information simultaneous price-setting game. Recall that $\mathcal{J}_j$ denotes the set of products produced by the firm that produces good j, e.g., the singleton $\{j\}$ in the special case of single-product firms or the full choice set $\mathcal{J}$ in the case of joint profit-maximization (perfect collusion).

Letting $\sigma_{jt} = \sigma_j(s_t, p_t)$ as shorthand, the first-order condition for the price of good j is

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

Following Bresnahan (1981, 1987), the first-order conditions for all firms can then be written in matrix form as

$$\sigma_t + \Delta_t (p_t - mc_t) = 0 \tag{A.1}$$

where $\sigma_t = (\sigma_{1t}, \ldots, \sigma_{Jt})'$ and the (j,k) 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 zero otherwise. Following BLP, the supply-side "inversion"for marginal cost is then

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

Assumption 7b will be satisfied if the matrix Δ_t is invertible. Observe that by simultaneous permutation of rows and columns (i.e., re-ordering of the goods to match the ownership structure), one can obtain from Δ_t a matrix that is block diagonal, with each block being a principle submatrix of $D_t(s_t, p_t)'$. Invertibility of this new matrix, and therefore of Δ_t, then follows from Lemma 3.

A.3 Quantity Setting

Consider the alternative model of simultaneous quantity-setting with complete information, again with Nash equilibrium as the solution concept. A fundamental requirement of this model is existence of an inverse demand function. Theorem 1 of Berry, Gandhi, and Haile (2012) ensures that under Assumption 2 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 $p_t = \rho(s_t, \delta_t)$.

If the inverse demand function is differentiable, equilibrium quantities are characterized by J first-order conditions of the form

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

Thus, Assumption 7b holds as long as the derivatives $\frac{\partial \rho_k}{\partial s_{jt}}$ exist. Lemma 3 ensures that the matrix $D_t(s_t, p_t)$ is invertible. So, given Assumption 6, the inverse function theorem ensures that derivatives of the inverse demand function exist and are equal to the elements of $D_t(s_t, p_t)^{-1}$, i.e.,

$$\frac{\partial \rho_k}{\partial s_{jt}} = \left[D_t(s_t, p_t)^{-1}\right]_{kj}.$$

B Identification with a Nonlinear Index

Here we discuss conditions under which the linearity of the indices δ_{jt} and κ_{jt} can be relaxed. One possible relaxation is to allow the nonlinear specification

$$\begin{aligned}\delta_{jt} &= g_j\left(x_{jt}\right)+\xi_{jt} \\ \kappa_{jt} &= h_j\left(\mathrm{w}_{jt}\right)+\omega_{jt}\end{aligned}$$

where g_j and h_j are unknown strictly increasing functions. This partial relaxation drops the requirement of linearity in the demand and cost shifters, while maintaining separability and monotonicity. Our constructive identification results using simultaneous equations (Theorems 6–8) can be extended to this setting by applying the results given in Theorem 1 or 2 in Berry and Haile (2011). The former result, first shown in Matzkin (2008), adds a restriction on the joint distribution of the structural errors (here, $(\xi_{1t},\ldots,\xi_{Jt},\omega_{1t},\ldots,\omega_{Jt})$) to the large support condition required in Theorem 6. The latter uses an alternative density restriction, but avoids any large support requirement. Because the extensions are direct applications of the results in Matzkin (2008) and Berry and Haile (2011), we omit a formal statement.

For the remainder of this appendix we consider an extension of the nonparametric IV approach, where the restrictions on the index can be relaxed further. We focus on the demand index and henceforth redefine

$$\delta_{jt}=\delta_j\left(x_{jt},\xi_{jt}\right)$$

with δ_j strictly increasing in its second argument. Relative to (2) this specification drops the assumptions of linearity, separability, and monotonicity in x_{jt}, retaining only the key requirement of strict monotonicity in ξ_{jt}.

Two lemmas will be useful. The first is taken from Berry, Gandhi, and Haile (2012) (their Lemma 2).

Lemma 4. *Under Assumptions 1 and 2, if δ and δ' are such that $I^+\equiv\left\{j:\delta_j'>\delta_j\right\}$ is nonempty, then $\sum_{j\in\mathcal{I}^+}\sigma_j\left(\delta',p\right)>\sum_{j\in\mathcal{I}^+}\sigma_j\left(\delta,p\right)$ for any p.*

Lemma 5. *Suppose Assumptions 1 and 2 hold. Then for all $j=1,\ldots,J$ and any s on the interior of Δ^J, $\sigma_j^{-1}\left(s,p\right)$ is strictly increasing in s_j.*

Proof. Arguing by contradiction, take $j=1$ without loss and suppose

$$s_1'>s_1,\ s_j'=s_j\forall j>1 \tag{B.1}$$

but

$$\delta_1'=\sigma_1^{-1}\left(s',p\right)\le\sigma_1^{-1}\left(s,p\right)=\delta_1.$$

Because probabilities sum to one, we must have $s_0'<s_0$, i.e.,

$$\sigma_0\left(\delta',p\right)<\sigma_0\left(\delta,p\right). \tag{B.2}$$

If the set $I^+ \equiv \{j : \delta'_j > \delta_j\}$ is nonempty, by Lemma 4 we have

$$\sum_{j \in I^+} \sigma_j(\delta', p) > \sum_{j \in I^+} \sigma_j(\delta, p).$$

Since $1 \notin I^+$ and $0 \notin I^+$, this contradicts the hypothesis that $s'_j = s_j$ for all $j > 1$. Thus, I^+ must be empty, i.e.,

$$\delta'_j \leq \delta_j \quad \forall j > 0.$$

But in that case Assumption 2 requires $\sigma_0(\delta', p) \geq \sigma_0(\delta, p)$, contradicting (B.2). □

Lemma 1 (which did not require a linear index) guarantees that for all j and t

$$\delta_j(x_{jt}, \xi_{jt}) = \sigma_j^{-1}(s_t, p_t)$$

for some functions σ_j^{-1}, $j = 1, \ldots, J$. Since $\delta_j(x_{jt}, \xi_{jt})$ is strictly increasing in ξ_{jt} we can then write

$$\begin{aligned} \xi_{jt} &= \delta_j^{-1}\left(\sigma_j^{-1}(s_t, p_t); x_{jt}\right) \\ &\equiv g_j(s_t, p_t, x_{jt}) \end{aligned}$$

for some unknown function g_j. Moreover, since $\delta_j(\cdot)$ is strictly increasing in ξ_{jt}, $\delta_j^{-1}(\cdot; x_{jt})$ is strictly increasing. So, by Lemma 5, g_j must be strictly increasing in s_{jt}. Inverting $g_j(\cdot)$, we obtain

$$\begin{aligned} s_{jt} &= g_j^{-1}\left(\xi_{jt}; s_{-jt}, p_t, x_{jt}\right) \\ &\equiv h_j\left(s_{-jt}, p_t, x_{jt}, \xi_{jt}\right) \end{aligned} \tag{B.3}$$

where s_{-jt} denotes $(s_{kt})_{k \neq j}$ and h_j is an unknown function that is strictly increasing in ξ_{jt}. The shares s_{jt} and s_{-jt} are bounded by definition, and we may assume without loss that p_t has been transformed to be bounded as well.

For arbitrary $j > 0$ consider the identification of the functions h_j in equation (B.3), which has the form of a nonseparable nonparametric regression function. Because x_{jt} is exogenous, we can condition on it and drop it from the notation, rewriting (B.3) as

$$s_{jt} = h_j\left(s_{-jt}, p_t, \xi_{jt}\right). \tag{B.4}$$

We will assume for simplicity that ξ_{jt} has an atomless marginal distribution. Then, without loss, we can normalize ξ_{jt} to have a standard uniform marginal distribution. We will also assume that (s_t, p_t) are continuously distributed conditional on (z_t, x_{-jt}) and let $f_{jsp}(s_{-jt}, p_t | z_t, x_{-jt})$ denote the conditional density of s_{-jt}, p_t. Let $f_{js}(s_{jt} | s_{-jt}, p_t, z_t, x_{-jt})$ denote the conditional density of s_{jt}.

Let ϵ_1 and ϵ_2 be some small positive constants. Let $\alpha_j(s_{-jt}, p_t)$ denote the set

$$\left\{s : f_{js}(s | s_{-jt}, p_t, z_t, x_{-jt}) \geq \epsilon_1 \ \forall (z_t, x_{-jt}) \ \text{such that} \ f_{jsp}(s_{-jt}, p_t | z_t, x_{-jt}) > 0\right\}.$$

For each j and $\tau \in (0,1)$ define $\mathcal{L}_j(\tau)$ as the convex hull of functions $m_j(\cdot,\cdot,\tau)$ that satisfy (a) for all (z_t, x_{-jt}), $\Pr(s_{jt} \leq m_j(s_{-jt}, p_t, \tau) | z_t, x_{-jt}) \in [\tau - \epsilon_2, \tau + \epsilon_2]$; and (b) for all (s_{-jt}, p_t), $m_j(s_{-jt}, p_t, \tau) \in \alpha_j(s_{-jt}, p_t)$. Consider the following instrumental variables conditions, from Chernozhukov and Hansen (2005, Appendix C).

Assumption 18. *$\xi_{jt} \perp\!\!\!\perp (z_t, x_{-jt}) \, \forall t$.*

Assumption 19. *For all $j = 1, \ldots, J$ and $\tau \in (0,1)$,*
(i) Let $\varepsilon_{jt} \equiv s_{jt} - h_j(s_{-jt}, p_t, \tau)$. For any bounded function $B_j(s_{-jt}, p_t, \tau) = m_j(s_{-jt}, p_t, \tau) - h_j(s_{-jt}, p_t, \tau)$ with $m_j(\cdot,\cdot,\tau) \in \mathcal{L}_j(\tau)$ and

$$\zeta_j(s_{-jt}, p_t, z_t, x_{-jt}, \tau) \equiv \int_0^1 f_{\varepsilon_j}(bB_j(s_{-jt}, p_t, \tau) | s_{-jt}, p_t, z_t, x_{-jt}) \; db > 0,$$

$E\left[B_j(s_{-jt}, p_t, \tau)\, \zeta_j(s_{-jt}, p_t, z_t, x_{-jt}, \tau) | z_t, x_{-jt}\right] = 0$ a.s. only if $B_j(s_{-jt}, p_t, \tau) = 0$ a.s.;
(ii) the conditional density $f_{\varepsilon_j}(e | s_{-jt}, p_t, z_t, x_{-jt})$ of ϵ_{jt} is continuous and bounded in e over $\mathbb{R}$ a.s.;
(iii) $h_j(s_{-jt}, p_t, \tau) \in \alpha_j(s_{-jt}, p_t)$ for all (s_{-jt}, p_t).

Assumption 18 strengthens the exclusion restriction of Assumption 3 to require full independence. Assumption 19 is a type of "bounded completeness" condition that replaces Assumption 4 in the text. It was used previously by Chernozhukov and Hansen (2005) to demonstrate nonparametric identification of quantile treatment effects.

Theorem 11. *Suppose Assumptions 1, 2, 18 and 19 hold. Then (i) ξ_{jt} is identified with probability one for all t, and (ii) the function $\sigma_j(\cdot)$ is identified on $\mathcal{X}$.*

Proof. Identification of $h_j(\cdot,\tau)$ for each $\tau \in (0,1)$ follows from Theorem 4 of Chernozhukov and Hansen (2005) after noting that for each value of $\xi_{jt} \in (0,1)$, the model (B.4) is equivalent to the model they consider. Parts (i) and (ii) then follow immediately, as in the proof of Theorem 1. □

Since j was arbitrary, this shows that the analog of Theorem 1 (i.e., identification of demand) can be obtained with the relaxed index structure.

C An Example: Multiple Endogenous Product Characteristics

As discussed in the text, extension of our results to the case of more than one endogenous product characteristic is straightforward provided that sufficient instruments are available. Here we point out that in some special cases, additional instruments may not be necessary, at least for identification of key features of demand such as own- and cross-price elasticities.

The important observation is that the exclusion conditions relied on in the text required independence (or mean independence) between instrument vector $x_t^{(1)}$ and the structural errors ξ_t conditional on $x_t^{(2)}$ (or, in some cases, between $\left(x_t^{(1)}, \mathrm{w}_t^{(1)}\right)$ and (ξ_t, ω_t) conditional on both $x_t^{(2)}$ and $\mathrm{w}_t^{(2)}$). This type of conditional independence is possible even when $\left(x_{jt}^{(1)}, x_{jt}^{(2)}, \xi_{jt}\right)$ are all jointly determined.

Consider a simple model of automobile product design. Suppose $x_{jt}^{(2)}$ consists of indicators for automobile class—e.g., luxury sedan, pickup truck, etc., $x_{jt}^{(1)}$ is a measure of size, and ξ_{jt} is an index representing unmeasured quality—e.g., leather seats, electronic features, and safety features.[32] Suppose further that

$$\begin{aligned} x_{jt}^{(1)} &= A\left(x_{jt}^{(2)}\right) + \epsilon_{jt}^{(1)} \\ \xi_{jt} &= B\left(x_{jt}^{(2)}\right) + \epsilon_{jt}^{(2)} \end{aligned}$$

where $\epsilon_{jt}^{(1)}, \epsilon_{jt}^{(2)}$ are independent of $x_{jt}^{(2)}$. An interpretation is that $A\left(x_{jt}^{(2)}\right)$ and $B\left(x_{jt}^{(2)}\right)$ are the "standard" size and unmeasured quality for an automobile in class $x_{jt}^{(2)}$, with $\epsilon_{jt}^{(1)}$ and $\epsilon_{jt}^{(2)}$ representing idiosyncratic deviations from the standard in each dimension arising through the design process. Here, size of the car $x_{jt}^{(1)}$, its class $x_{jt}^{(2)}$, and its unobserved quality ξ_{jt} are all correlated. However, conditional on $x_{jt}^{(2)}$, $x_{jt}^{(1)}$ is independent of ξ_{jt}.

Thus, in this example, identification of demand (conditional on $x_t^{(2)}$) follows from Theorem 1 without modification. Note, however, that the effects of an exogenous change in $x_t^{(2)}$ on demand would not be identified without additional structure such as that provided by the example.

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.

Andrews, D. T. K. (2011): "Examples of L^2-Complete and Boundedly Complete Distributions," Discussion paper, Yale University.

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

[32]It should be clear that additional observables that are exogenous conditional on $x_t^{(2)}$ are easily accommodated.

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.

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

Berry, S., and P. Jia (2010): "Tracing the Woes: An Empirical Analysis of the Airline Industry," *American Economic Journals: Microeconomics*.

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., A. Gandhi, and P. A. Haile (2012): "Connected Substitutes and Invertibility of Demand," Discussion paper, Yale University.

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

——— (2010): "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," Discussion Paper 1718, Cowles Foundation at Yale.

——— (2011): "Identification in a Class of Nonparametric Simultaneous Equations Models," Discussion Paper 1787R, Cowles Foundation at Yale.

Block, H., and J. Marschak (1960): "Random Orderings and Stochastic Theories of Responses," in *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, ed. by I. Olkin, S. Ghurye, W. Hoeffding, W. G. Mado, and H. B. Mann. Stanford University Press.

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 American Automobile Industry," *Journal of Econometrics*, 17, 201–227.

——— (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.

Bundorf, K., J. Levin, and N. Mahoney (12): "Pricing and Welfare in Health Plan Choice," *American Economic Review, forthcoming.*

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.

Chernozhukov, V., and C. Hansen (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73(1), 245–261.

Chesher, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405–1441.

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. (2011): "Upstream Innovation and Product Variety in the United States Home PC Market," Discussion paper, Hebrew University.

Fan, Y. (2011): "Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market," Discussion paper, University of Michigan.

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

——— (2011): "Identifying Demand with Multi-Dimensional Unobservables: A Random Functions Approach," Discussion paper, University of Wisconsin.

Gandhi, A., K.-I. Kim, and A. Petrin (2011): "Identification and Estimation in Discrete Choice Demand Models when Endogenous Variables Interact with the Error," Discussion paper, University of Wisconsin-Madison.

Gandhi, A., and A. Nevo (2011): "Flexible Estimation of Random Coefficient Discrete Choice Model Using Aggregate Data," Discussion paper, University of Wisconsin.

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

Gentzkow, M. (2004): "Valuing New Goods in a Model with Complementarity: Online Newspapers," Working paper, Chicago GSB.

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

Goeree, M. S. (2008): "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 76, 1017–1074.

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.

Goolsbee, A., and A. Petrin (2004): "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 72, 351–381.

Gordon, B., and W. R. Hartmann (2010): "Structural Equilibrium Analysis of Political Advertising," Discussion paper, Columbia University.

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

Hastings, J., T. Kane, and D. Staiger (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. (2009): "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, 99, 393–430.

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.

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.

Imbens, G. W., and W. K. Newey (2009): "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77(5), 1481–1512.

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

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. (1977): "The Structure of Random Utility Models," *Theory and Decision*, 8, 229–254.

——— (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.

Marschak, J. (1960): "Binary Choice Constraints on Random Utility Indicators," in *Mathematical Methods in the Social Sciences*, ed. by K. Arrow, pp. 312–329. Stanford University Press.

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.

——— (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.

——— (2010): "Estimation of Nonparametric Models with Simultaneity," Discussion paper, UCLA.

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

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.

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.

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

Souza-Rodrigues, E. (2011): "Nonparametric Estimator of a Generalized Regression Model with Group Effects," Working paper, Yale University.

Thompson, T. S. (1989): "Identification of Semiparametric Discrete Choice Models," Discussion paper no. 249, Center for Economics Research, Department of Economics, University of Minnesota.

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