IDENTIFICATION IN DIFFERENTIATED PRODUCTS MARKETS USING MARKET LEVEL DATA

Steven T. Berry
Philip A. Haile

WORKING PAPER 15641

NBER WORKING PAPER SERIES

IDENTIFICATION IN DIFFERENTIATED PRODUCTS MARKETS USING MARKET LEVEL DATA

Steven T. Berry
Philip A. Haile

Working Paper 15641
http://www.nber.org/papers/w15641

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2010

Earlier drafts were circulated under the title "Identification of Discrete Choice Demand From Market Level Data." We had helpful early conversations on this topic with Rosa Matzkin and Yuichi Kitamura. We also had valuable feedback from Liran Einav, Jon Levin, Tim Bresnahan, and participants in the 2008 Game Theory World Congress, the 2008 LAMES, the 2009 FESAMES, 2009 SETA Conference, the Brown Conference on the Econometrics of Demand, and seminars at Boston College, CREST, Penn, Stanford, USC, UCSD, Virginia, and Wisconsin. Financial support from the NSF is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Identification in Differentiated Products Markets Using Market Level Data
Steven T. Berry and Philip A. Haile
NBER Working Paper No. 15641
January 2010
JEL No. C35,L0

ABSTRACT

We consider nonparametric identification in models of differentiated products markets, using only market level observables. On the demand side we consider a non-parametric random utility model nesting random coefficients discrete choice models widely used in applied work. We allow for product/market-specific unobservables, endogenous product characteristics e.g., prices), and high-dimensional taste shocks with arbitrary correlation and heteroskedasticity. On the supply side we specify marginal costs nonparametrically, allow for unobserved firm heterogeneity, and nest a variety of equilibrium oligopoly models. We pursue two approaches to identification. One relies on instrumental variables conditions used previously to demonstrate identification in a nonparametric regression framework. With this approach we can show identification of the demand side without reference to a particular supply model. Adding the supply side allows identification of firms' marginal costs as well. Our second approach, more closely linked to classical identification arguments for supply and demand models, employs a change of variables approach. This leads to constructive identification results relying on exclusion and support conditions. Our results lead to a testable restriction that provides the first general formalization of Bresnahan's (1982) intuition for empirically discriminating between alternative models of oligopoly competition.

Steven T. Berry
Yale University Department of Economics
Box 208264
37 Hillhouse Avenue
New Haven, CT 06520-8264
and NBER
steven.berry@yale.edu

Philip A. Haile
Department of Economics
Yale University
37 Hillhouse Avenue
P.O. Box 208264
New Haven, CT 06520
and NBER
philip.haile@yale.edu

1 Introduction

Models of discrete choice between differentiated products play a central role in the modern empirical literature in industrial organization (IO) and are used in a wide range of other applied fields of economics.[1] In many applications, the discrete choice demand model is combined with an oligopoly model of supply. 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 counterfactuals. However, an important question is whether parametric specifications and distributional assumptions play a more fundamental role in determining what is learned from the data. In particular, are such assumptions essential for identification?

In this paper we examine the nonparametric identifiability of differentiated products models that are in the spirit of Berry, Levinsohn, and Pakes (1995) (henceforth, "BLP") and a large applied literature that has followed. We consider identification of demand alone, identification of marginal costs, and discrimination between different models of oligopoly competition.

On the demand side, the models motivating our work incorporate two essential features. One is rich heterogeneity in preferences, which allows for flexible demand substitution pat-

[1] Applications include studies of the sources of market power (e.g., Berry, Levinsohn, and Pakes (1995), Nevo (2001)), welfare gains from new goods or technologies (e.g., Petrin (2002), Eizenberg (2008)), mergers (e.g., Nevo (2000), Capps, Dranove, and Satterthwaite (2003)), network effects (e.g., Rysman (2004), Nair, Chintagunta, and Dube (2004)), product promotions (e.g., Chintagunta and Honoré (1996), Allenby and Rossi (1999)), environmental policy (e.g., Goldberg (1998)), vertical contracting (e.g., Villas-Boas (2007), Ho (2007)), market structure and product quality (e.g., Fan (2008)), media bias (e.g., Gentzkow and Shapiro (2009)), asymmetric information and insurance (e.g., Cardon and Hendel (2001), Bundorf, Levin, and Mahoney (2008), Lustig (2008)), trade policy (e.g., Goldberg (1995), Berry, Levinsohn, and Pakes (1999), Goldberg and Verboven (2001)), residential sorting (e.g., Bayer, Ferreira, and McMillan (2007)), and school choice (e.g., Hastings, Staiger, and Kane (2007)).

terns (e.g., cross-price elasticities).[2] The second is the presence of product/market-level unobservables.[3] These unobservables give rise to the endogeneity of prices, and only by explicitly modeling them can one account simultaneously for endogeneity and heterogeneity in preferences for choice characteristics. Surprisingly, this combination of factors has not been treated in the prior literature on identification. Indeed, although there is a large literature on identification of discrete choice models,[4] 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 applied literature employs empirical models derived from equilibrium conditions for multi-product oligopolists, building on early insights of Rosse (1970) and Bresnahan (1981). Following BLP, recent work typically allows for latent cost shocks and unobserved heterogeneity in cost functions, but employs a parametric specification of costs.

We consider identification within nonparametric generalizations of the demand-side and supply-side models used in the applied literature. We focus on the common situation in which 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. Individual choices and firms' costs are not observed.[5]

We begin with a nonparametric generalization of standard random coefficients discrete choice models. Our model incorporates an important index restriction on the way product/market-

[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).

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

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

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

specific unobservables enter preferences, but is otherwise flexible. It allows for market/choice-specific unobservables, endogenous choice characteristics, unknown heteroskedasticity, and taste shocks with arbitrary dimension and correlation. We consider identification of demand as well as full identification of the joint distribution of consumers' conditional indirect utilities, the latter enabling characterization of standard aggregate welfare measures. Identification of demand naturally requires instruments for prices (or other endogenous choice characteristics), and we show that standard nonparametric instrumental variables conditions (Newey and Powell (2003)) suffice. This result demonstrates that the essential requirement for identification of demand in this type of model is identical to that for regression models: the availability of instruments. Further, this result can be extended to full identification of the random utility model using standard arguments under additional quasilinearity and support conditions.

Given identification of demand, we consider nonparametric identification of each firm's marginal cost function and cost shocks, again relying on an index restriction and on nonparametric instrumental variables conditions. These results provide a nonparametric foundation for a large body of applied work that estimates marginal costs in order to address positive and normative questions concerning imperfectly competitive markets.

Although these are strong positive results, nonparametric instrumental variables conditions themselves can be difficult to interpret or verify. This is one reason we consider a second approach to identification, this time making simultaneous use of the demand model and a partially specified model of oligopoly competition. We show that the resulting system of "supply and demand" equations can be "inverted," leading to identification of demand and of the latent shocks to marginal costs. If we further commit to a particular oligopoly model (e.g., Nash equilibrium in prices) we also recover each firm's marginal costs. This approach enables us to offer constructive identification arguments relying on traditional exclusion and support conditions, but under somewhat stronger restrictions on the model.

Finally, we show that (using either identification approach) our results lead to testable restrictions that can distinguish between alternative models of oligopoly competition. This result offers the first general formalization of Bresnahan's (1982) intuition for empirically discriminating between alternative "oligopoly solution concepts."

Together these results provide a positive message regarding the faith we may have in a growing body of applied work on differentiated products markets allowing for rich consumer and firm heterogeneity, choice-specific unobservables, and endogeneity. Such models are identified without parametric or distributional assumptions under the same sorts of conditions that yield identification of simpler and more familiar models. This positive message is not without qualification: in addition to the index restrictions, our results require instruments for price that are excluded from the demand system, as well as demand shifters that are excluded from firm costs. Of course, it should not be surprising that identification in an environment with limited dependent variables and endogeneity requires both some structure and adequate exogenous variation. Our results shed light on key assumptions and essential sources of variation one should look for in applications.

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

2 Related Literature

Our work relates to several literatures, including a large body of work on identification of discrete choice models. 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) studied a linear random coefficients binary choice model. Briesch, Chintagunta, and Matzkin (2005) consider multinomial choice, allowing some generalization of a linear random coefficients model. Our work relaxes functional form and distributional assumptions relied on in this earlier work, incorporates market/choice-specific unobservables, and allows for endogeneity.

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

Compared to these papers, we relax functional form restrictions and, more fundamental, add the important distinction between market/choice-specific unobservables and individual heterogeneity in preferences. This distinction allows the model to define comparative statics that account for both heteroskedasticity (heterogeneity in tastes for characteristics) and endogeneity.[6] 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

[6] 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).

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.

Blundell and Powell (2004), Matzkin (2004), and Hoderlein (2008) have considered binary choice with endogeneity in semiparametric triangular models, leading to the applicability of control function methods or the related idea of "unobserved instruments" (see also Petrin and Train (2009), Altonji and Matzkin (2005), Gautier and Kitamura (2007), and Fox and Gandhi (2009)). However, standard models of oligopoly pricing in differentiated products markets imply that each equilibrium price depends on the entire vector of demand shocks (and typically the vector of cost shocks as well). This rules out a triangular structure. Nonetheless, our "change of variables" approach uses a related strategy of inverting a multiproduct supply and demand system to recover the entire vector of shocks to costs and demand. This can be interpreted as a generalization of the control function approach.

On the supply side, Rosse (1970) introduced the idea of using first-order conditions for imperfectly competitive firms to infer their marginal costs from prices and demand parameters.[7] Our approach to identification of marginal costs is a nonparametric extension of that idea.

Our insights regarding discrimination between alternative oligopoly models are closely related to ideas from the early empirical IO literature on conjectural variations models. 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, and have required shifters of aggregate demand.

Our change of variables approach, which exploits the simultaneous determination of prices and market shares, has links to the prior literature on nonparametric identification of simul-

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

taneous equations models (e.g., Brown (1983), Roehrig (1988), Matzkin (2005), and Matzkin (2008)). A standard strategy in this literature is to relate the joint density of latent variables to that of the observables using restrictions from theory and a standard change of variables. A complication, emphasized in Benkard and Berry (2006), is that the change of variables involves the Jacobian of the transformation. This introduces substantial challenges and has limited the set of models for which identifiability has been shown using the change of variables approach. However, in our context the same index restriction that enables us to use the nonparametric instrumental variables strategy permits us to use a new change of variables argument to obtain a constructive proof of identification. Here our work is closely related to that of Matzkin (2005, 2008), who has explored identification in a variety of nonparametric simultaneous equations models. Although she does not explicitly address discrete choice models, for our change of variable argument we transform our model to a form equivalent to one she considers. This transformation maps our index restriction to a separability condition whose advantages she emphasizes in a variety of other contexts. Even starting from the transformed model, however, our assumptions and proof differ from hers in important ways.[8] Our strategy and results may therefore complement those in Matzkin (2008) for other applications of simultaneous equations models.

For an important preliminary result, we rely heavily on insights in Gandhi (2008), which recently showed how to extend a key invertibility result of Berry (1994) and Berry and Pakes (2007) to a more general class of discrete choice demand models. We reinterpret Gandhi's key assumption graphically as our "connected substitutes" condition, requiring that for every pair of products $\{j,k\}$ there be some path of local substitution linking j to k. Although Gandhi (2008) focused on invertibility of demand, we show that the same connected substitutes condition plays an important role in ensuring the invertibility of the oligopoly supply side.

Turning to other recent unpublished papers, Berry and Haile (2009b) explores the iden-

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

tification of discrete choice models in the case of "micro data" relying in part on ideas similar to those used here. The distinction between "market data" and "micro data" has been emphasized in the recent industrial organization literature (e.g., Berry, Levinsohn, and Pakes (2004)), but not the econometrics literature on identification. A key insight in Berry and Haile (2009b) is that within a market all market/choice-specific unobservables are held fixed. One can therefore learn a great deal about the distribution of utilities from "variation in choice sets" created by within-market heterogeneity in consumer/choice-specific covariates—variation that is not confounded by variation in the market/choice-specific unobservables. That strategy is exploited throughout Berry and Haile (2009b), but cannot be applied to market level data.[9] In Berry and Haile (2009a) we have explored related ideas in the context of a "generalized regression model" (Han (1987)), which nests the binary choice model. For that class of models, the index restriction we require throughout the present paper can be dropped. Unfortunately, many applications fall outside the binary choice setting.

Concurrent work by Fox and Gandhi (2009) explores identifiability of several related models, including a flexible model of polychotomous choice in which consumer types are themselves multinomial and the conditional indirect utility functions are analytic. They do not consider our case with market-level data and endogenous prices set in a (non-triangular) system of equations.[10] A recent working paper by Chiappori and Komunjer (2009) considers a related change of variables approach in a "micro data" context.

[9] Berry and Haile (2009b) includes an example in which what appears to be a "market data" environment is actually isomorphic to the "micro data" environment. In that example one has a continuum of observations for which choice-specific unobservables are held fixed while observables vary. In general this is not the case.

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

3 Demand Model

3.1 Consumers, Products and Markets

Each consumer i in market t chooses a single good from a set $\mathcal{J}_t$. We will use the terms "good," "product," and "choice" interchangeably. The term "market" is synonymous with the choice set. In practice, markets will typically be defined geographically and/or temporally. The choice set always includes the option not to purchase, i.e., to choose the "outside good," which we index as choice $j = 0$. We denote the number of "inside goods" by $J_t = |\mathcal{J}_t| - 1$.[11]

Each inside good/market has observable (to us) characteristics $x_{jt} \in \mathbb{R}^{K_x}$ and price $p_{jt} \in \mathbb{R}$. We treat x_{jt} and p_{jt} differently because we will allow p_{jt} to be endogenous. The restriction to a single endogenous characteristic reflects the usual practice, but is not essential.[12] We allow x_{jt} to include components that vary only with the market t, only with the product j (these could be product dummies), or both. Unobservables at the level of the product and market are represented by an index $\xi_{jt} \in \mathbb{R}$. In applications this is typically motivated by the presence of unobserved product characteristics and/or unobserved variation in tastes across markets. A market (choice set) t is thus characterized by $\left(\mathcal{J}_t, \left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}\right)$. We let $\chi = \text{supp}\left(x_{jt}, p_{jt}, \xi_{jt}\right)$ and $\chi^{\mathcal{J}_t}$ =supp$\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}$.

3.2 Preferences

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

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

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

discuss restrictions on $\mathcal{U}$ below.

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

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

where u is measurable in ω_{it}, and $u(\cdot, \cdot, \cdot, \omega) \in \mathcal{U}$ for all $\omega \in \Omega$. Thus, conditional indirect utilities are represented with a random function $u : \chi \times \Omega \to \mathbb{R}$.[13]

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

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

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

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

$$\left(\alpha\left(\omega_{it}\right), \beta^{(1)}\left(\omega_{it}\right), \ldots, \beta^{(K_x)}\left(\omega_{it}\right), \epsilon_1\left(\omega_{it}\right), \ldots, \epsilon_J\left(\omega_{it}\right)\right).^{14}$$

This structure permits $(\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ *to have an arbitrary joint distribution but is more restrictive than necessary. For example, specifying* $\epsilon_{ijt} = \epsilon_j\left(x_{jt}, p_{jt}, \omega_{it}\right)$ *would allow richer preference heterogeneity/heteroskedasticity. If we relax the model further by specifying*

[13] See, e.g., Gikhman and Skorokhod (1980).

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

$\epsilon_{ijt} = \epsilon_j\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right)$, *the terms* $x_{jt}\beta_{it} - \alpha_{it}p_{jt} + \xi_{jt}$ *in (2) become redundant and we obtain our original model* $u\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right) = \epsilon\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right)$.

Note that there is no market subscript t on the probability measure $\mathbb{P}$. This reflects our assumption that ξ_{jt} captures all unobserved heterogeneity at the market and/or product level. This is standard in the literature but is an important restriction.[15] Aside from this restriction, however, our representation of preferences is so far fully general. For example, it allows arbitrary correlation of consumer-specific tastes for different goods or characteristics. Because x_{jt} can include product dummies, it allows marginal utilities of characteristics to differ arbitrarily across products. It also allows arbitrary heteroskedasticity in utilities across different products, or in utilities for a given product j as its characteristics $\left(x_{jt}, \xi_{jt}\right)$ vary.

However, we will rely on two restrictions on preferences throughout the paper. First, for simplicity we will assume that for all functions $\tilde{u}_{it} \in \mathcal{U}$, any $\mathcal{J}_t$, any distinct $k, \ell \in \mathcal{J}_t$, and any $\left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j\in\{k,\ell\}}$, the random differences $\tilde{u}_{it}\left(x_{kt}, p_{kt}, \xi_{kt}\right) - \tilde{u}_{it}\left(x_{\ell t}, p_{,\ell t}, \xi_{\ell t}\right)$ are continuously distributed with convex support. This simplifies the analysis by enabling us to ignore ties and choice probabilities that are invariant to a strict (stochastic) increase in v_{ijt} for some j. The second restriction is more significant, imposing an index restriction on the way market/choice-specific unobservables enter the random utility function. To state this condition, partition x_{jt} as $\left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, with $x_{jt}^{(1)} \in \mathbb{R}$, and define the index

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

Assumption 1a. For all $\tilde{u}_{it} \in \mathcal{U}$, $\tilde{u}_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right) = \mu_{it}\left(\delta_{jt}, x_{jt}^{(2)}, p_{jt}\right)$ for some function μ_{it} that is strictly increasing in its first argument.

This assumption limits attention to random utility functions admitting representations

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

of the form

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

There are two parts to this restriction. The first is a restriction to a "vertical" unobservable ξ_{jt}; i.e., all else equal, an increase in ξ_{jt} makes product j more attractive to all consumers. This is standard in the applied literature and is a property we rely on to allow recovery of each ξ_{jt}.[16] The second is a linear index restriction requiring perfect substitutability between ξ_{jt} and $x_{jt}^{(1)}$ inside the function u.[17] As already mentioned, this index restriction plays an important role in both of our identification approaches. However, we show in Appendix B that strengthening the instrumental variables requirements used for our nonparametric IV approach would allow us to replace the linear indices $x_{jt}^{(1)} + \xi_{jt}$ with nonlinear indices $\delta_j\left(x_{jt}^{(1)}, \xi_{jt}\right)$ that are strictly monotonic in ξ_{jt}.

With this structure on preferences we will show identification of demand, identification of marginal costs, and the falsifiability of a given oligopoly model. To obtain results that also allow characterization of standard welfare measures, we will strengthen Assumption 1a to require quasilinearity in price.[18]

Assumption 1b. For all $\tilde{u}_{it} \in \mathcal{U}$ there is a monotonic function Γ_{it} such that $\Gamma_{it}\left(\tilde{u}_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right) = \mu_{it}\left(\delta_{jt}, x_{jt}^{(2)}\right) - p_{jt}$ for some function μ_{it} that is strictly increasing in its first argument.

Assumption 1b differs from Assumption 1a in requiring conditional indirect utilities with

[16] Athey and Imbens (2007) and Berry and Haile (2009b) discuss testable implications in the case of micro data.

[17] In the linear random coefficients model of Example 1, Assumption 1a holds if either (a) one covariate enters with a fixed coefficient (common in practice) or (b) $u_{ijt} = x_{jt}^{(1)}\beta_{it}^{(1)} + \xi_{jt}\beta_{it}^{(1)} + x_{jt}^{(2)}\beta_{it}^{(2)} - \alpha_{it}p_{jt} + \epsilon_{ijt}$, with $\beta_{it}^{(1)} > 0$. Of course, we do not require the linear random coefficients structure.

[18] We assume quasilinearity in price because this is the natural unit for measuring compensating variation. However, quasilinearity in any element of $x_{jt}^{(2)}$ (or in δ_{jt}) would also suffice.

quasilinear representations[19]

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

Quasilinearity plays two roles for us. First, it allows the model to define standard aggregate welfare measures. Second, it provides a mapping between the observable units of choice probabilities and the latent units of utilities. This is a standard strategy in the literature on discrete choice. Note, however, that we do not require independence between p_{jt} and the random variable $\mu\left(\delta_{jt}, x_{jt}^{(2)}, \omega_{it}\right)$; in fact, this is ruled out by standard models of supply. However, the restriction that $\mathbb{P}$ does not vary across markets does imply that $\mu\left(\delta_{jt}, x_{jt}^{(2)}, \omega_{it}\right)$ is independent of p_{jt} conditional on $\left(\delta_{jt}, x_{jt}^{(2)}\right)$.[20] To relate this to more familiar models, observe that in the linear random coefficients model of Example 1, this conditional independence holds if $(\beta_{it}, \epsilon_{it})$ are assumed independent of p_{jt}.

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

[19] As long as v_{ijt} is strictly decreasing in p_{jt} with probablity one, (4) is equivalent to a spefication allowing a random coefficient on price: the scale of each consumer's utility function can be normalized by her price coefficient without loss.

[20] Recall that each market is t is defined by $\left(\mathcal{J}_t, \{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}_t}\right)$.

3.3 Market Shares

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

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

We denote consumer i's choice by

$$y_{it} = \arg\max_{j \in \mathcal{J}_t} \; u\left(\delta_{jt}, x_{jt}^{(2)}, p_{jt}, \omega_{it}\right).$$

Given $\mathcal{J}_t$, market shares (choice probabilities) are given by

$$\begin{aligned} s_{jt} &= E_{\mathbb{P}}\left[1\left\{y_{it} = j\right\}\right] \\ &= E_{\mathbb{P}}\left[1\left\{u\left(\delta_{jt}, x_{jt}^{(2)}, p_{jt}, \omega_{it}\right) > u\left(\delta_{kt}, x_{kt}^{(2)}, p_{kt}, \omega_{it}\right) \; \forall k \in \mathcal{J}_t - \{j\}\right\}\right] \\ &\equiv s_j\left(\{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}_t}\right). \end{aligned} \tag{6}$$

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

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

Market shares are positive for all goods at *any* $\{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}_t}$ in models for which the support of $\{v_{ijt}\}_{k \in \mathcal{J}_t}$ is always $\mathbb{R}^{J_t}$. In a parametric context this includes multinomial probit or logit models, for example. Assumption 2 of course requires only that market shares be positive at *equilibrium* values of $\{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}_t}$.[21]

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

3.4 Observables and Primitives of Interest

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

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

Identification of demand is sufficient for many purposes motivating demand estimation. However, we also consider identification of the joint distribution of indirect utilities conditional on the choice set $(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$. We refer to this as *full identification of the random utility model*. These conditional distributions are the primitives determining all quantities defined by our random utility model. Under the quasilinear structure of Assumption 1b this includes standard measures of aggregate welfare.[22]

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

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

and

$$v_{ijt} = \mu_j\left(\delta_{jt}, \omega_{it}\right) - p_{jt} \tag{8}$$

[22] The primitives of our model do not determine individual utilities. Thus, for example, it cannot be used to characterize Pareto improvements. As usual, assuming quasilinearity will allow us to characterize potential Pareto improvements. Parametric random utility models allow tracking of individual utilities (even with market level data!) by associating realizations of random preference parameters with individual consumers.

(with j subscripts on the functions u and μ) to represent, respectively, (3) and (4) above, since the utility functions will generally be evaluated at different $x_{jt}^{(2)}$ for each j.[23] We will work with these two representations of preferences in what follows.

4 Connected Substitutes

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

Definition 1. *Product k substitutes to product ℓ at* $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}$ if, for any variable w_{kt} such that v_{ikt} is strictly stochastically increasing in w_{kt}, $s_\ell\left(\{x_{jt}, p_{jt}, \xi_{jt}\}_{j\in\mathcal{J}}\right)$ rises when w_{kt} falls.[24]

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

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

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

---

[23] Because $x_t^{(2)}$ may include product dummies, the functions u_j and μ_j may vary arbitrarily with j.

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

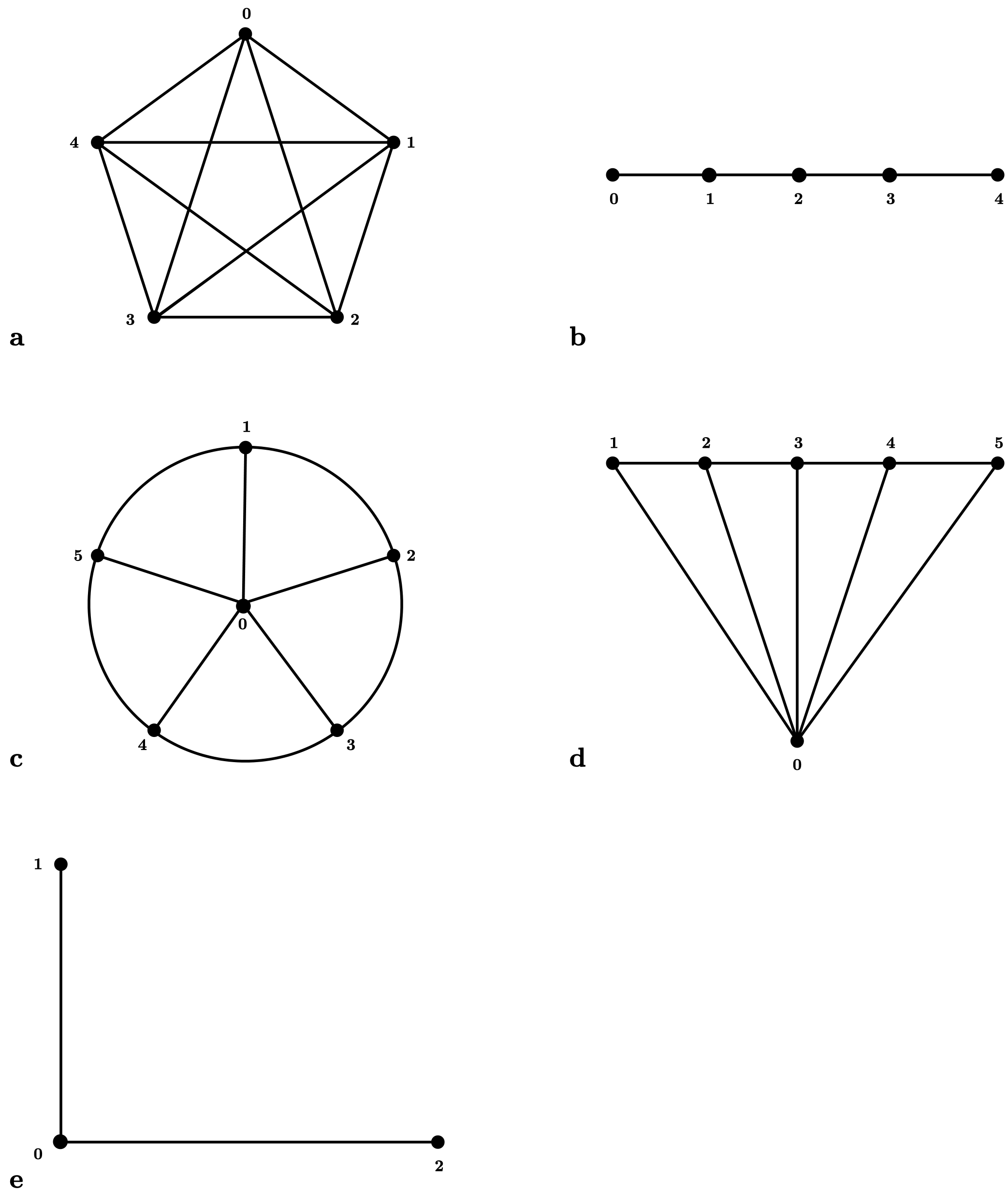

Figure 1: Graphs of $\Sigma(\mathcal{J})$ at generic $\{x_{jt}, p_{jt}, \xi_{jt}\}_{j \in \mathcal{J}}$ for standard models. All edges are bi-directional, so for simplicity we show undirected graphs. Panel a: multinomial logit, multinomial probit, mixed logit, etc.; Panel b: pure vertical models e.g.,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

The directed graph of $\Sigma(\mathcal{J})$ has nodes (vertices) representing each product and an edge from product k to product ℓ whenever product k substitutes to product ℓ.[25] Formally, the connected substitutes condition requires that for any distinct products j and j' there be a path of substitution, possibly indirect, from j to j'. To describe the key economic implication, interpret w_{kt} in Definition 1 as the "quality" of product k. Then Assumption 3 implies that, all else equal, if "quality" rises (falls) for every product in some strict subset $I \subset \mathcal{J}$ but for no product outside I, the total market share of all products outside I falls (rises).[26] This is a very natural property for an environment with unit demands.

It is also easy to verify that the connected substitutes property itself holds in standard models. In fact, the usual random utility discrete choice models (e.g., multinomial probit, logit, mixed logit, etc.) imply that every product substitutes *directly* to every other product, a strong sufficient condition for connected substitutes. Figure 1 illustrates, showing the graphs of $\Sigma(\mathcal{J})$ for a variety of models. As panel e illustrates, even a market comprised of independent goods satisfies this condition as long as each product substitutes to and from the outside good.

5 Identification with Nonparametric IV Conditions

5.1 Identification of Demand

Let $x_t = (x_{1t}, \ldots, x_{Jt})$, $p_t = (p_{1t}, \ldots, p_{Jt})$, and $\delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$. Under Assumption 1a, for any vector δ_t, market shares are given by

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

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

[26] This is demonstrated formally as Lemma 4 in Appendix A. An analogous implication follows if we interpret $-\mathrm{w}_{jt}$ as the price of good j.

Using the connected substitutes assumption, we can follow the argument in Theorem 2 of Gandhi (2008) to show the following lemma, which generalizes well-known invertibility results for linear discrete choice models in Berry (1994) and Berry and Pakes (2007).[27]

Lemma 1. *Consider any price vector p and any market share vector $s = (s_1, \ldots, s_J)'$ on the interior of $\triangle^J$. Under Assumptions 1a, 2, and 3 there is at most one vector δ such that $\sigma_j(\delta, p) = s_j \; \forall \; j$.*

Proof. See Appendix A.

With this result we can write

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

or

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

To state the instrumental variables conditions, recall that $\tilde{z}_t$ represents instruments for p_t excluded from the determinants of $\{v_{ijt}\}_{j \in \mathcal{J}}$. Standard excluded instruments include cost shifters (e.g., input prices) or proxies for cost shifters such as prices of the same good in other markets (Hausman (1996), Nevo (2001)). With the exogenous conditioning variables $(\tilde{z}_t, x_t)$, we take the following instrumental variables assumptions from Newey and Powell (2003).

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

[27] See Corollary 1 in Appendix A for the equivalence between Gandhi's condition and our connected substitutes condition. Berry (1994) and Berry and Pakes (2007) show existence and uniqueness of an inverse choice probability in models with an additively separable δ_{jt}. Gandhi (2008) relaxes the separability requirement. Our lemma addresses only uniqueness conditional on existence. Under our maintained assumption that the model is correctly specified, given any observed choice probability vector, there must exist a vector $(\delta_1, \ldots, \delta_J)$ that rationalizes it. Gandhi (2008) provides conditions guaranteeing that an inverse exists for every choice probability vector in $\triangle^J$. Our uniqueness result differs from his only slightly, mainly in recognizing that the argument applies to a somewhat more general model of preferences. For a class of single-agent dynamic discrete choice models similar to our model with an additive δ_{jt}, Hotz and Miller (1993) prove the uniqueness of the inverse of the share function and sketch a proof of existence of the inverse that is correct for the binary case.

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

Assumption 4 is a standard exclusion restriction, requiring mean independence between the instruments and the structural error ξ_{jt}. Note that setting $E[\xi_{jt}|\tilde{z}_t, x_t]$ equal to zero rather than another constant provides the required normalization of the location of ξ_{jt}. Assumption 5 is a "completeness" condition, which is the nonparametric analog of the standard rank condition for linear models. This condition requires that the instruments move the endogenous variables (s_t, p_t) sufficiently to ensure that any function of these variables can be distinguished from other functions through the exogenous variation in the instruments. See Newey and Powell (2003) (and references therein) and Severini and Tripathi (2006) for helpful discussion and examples.[28]

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

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

Proof. For any j, rewriting (11) and taking expectations conditional on $\tilde{z}_t, x_t$, we obtain

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

so that by Assumption 4,

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

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

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

[28] If we assumed bounded support for ξ_{jt} and x_{jt} we could replace the completeness condition with bounded completeness.

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

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

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

The proof is very similar to that given by Newey and Powell (2003) in the context of nonparametric regression. A difference is that we have, in addition to the nonparametric function $\sigma_j^{-1}(s_t, p_t)$, the additive x_{jt}, which drops out of the proof and becomes available as as one of the $2J$ instruments for the $2J$ endogenous variables (s_t, p_t).

If we strengthen the instrumental variables assumptions 4 and 5 as in Chernozhukov and Hansen (2005), we can relax the linear index structure, allowing

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

where each δ_j is any function that is strictly increasing in its second argument. We discuss this extension in Appendix B.

5.2 Full Identification of the Random Utility Model

We consider full identification of the random utility model under the quasilinear specification of preferences in Assumption 1b. We will also make a large support assumption:

Assumption 6. supp $p_t | \left\{ x_{jt}, \xi_{jt} \right\}_{j \in \mathcal{J}} = \mathbb{R}^J$.

This type of support condition is strong but also standard in the literature because it provides a natural benchmark for evaluating identification under ideal conditions on observables. It is intuitive that in order to trace out the full CDF of the random part of a random utility model, extreme values of observables will be needed.[29] As the proof of the following result makes clear, we use the large support condition only for this role; in particular, we do not use the common "identification at infinity" argument that takes observables for all but one choice to extreme values in order to reduce a multinomial choice problem to a binary choice problem. The argument here makes clear that if the support condition fails (the support of p_t excludes tail values), the implication will be that the joint distributions of $(v_{i1t}, \ldots, v_{iJt})$ will be unknown at its tail values.

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

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

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

By Theorem 1 each ξ_{jt} is identified, so each δ_{jt} can be treated as known. Then, since p_{jt} is independent of $\mu\left(\delta_{jt}, \omega_{it}\right)$ conditional on δ_{jt}, Assumption 6 ensures identification of the joint distribution of

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

for any $(\delta_{1t}, \ldots, \delta_{Jt})$. Since $v_{ijt} = \mu_j\left(\delta_{jt}, \omega_{it}\right) - p_{jt}$, the result follows. □

5.3 Adding A Supply Side

If we are willing to add a specification of the supply model, we can obtain identification of firms' marginal costs as well. Here we can return to the less restrictive specification of preferences in Assumption 1a. Our approach generalizes arguments from the parametric literature on the estimation of static oligopoly models, which utilize first-order conditions for

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

firms to solve for marginal costs in terms of demand parameters. Using first-order conditions requires that the market share function $\sigma_j(\delta_t, p_t)$ be differentiable with respect to prices, and we will assume this directly.

Assumption 7. $\sigma_j(\delta_t, p_t)$ is continuously differentiable with respect to p_k $\forall j, k \in \mathcal{J}$.

We consider a nonparametric specification of costs, but require sufficient structure to ensure that behavior is characterized by first-order conditions that can be inverted to solve for the unobserved cost shocks. As with the demand model, the most restrictive assumption we require is an index restriction on how these shocks enter. In particular, we assume the marginal cost associated with product j depends on its output quantity $q_{jt} = M_t \sigma_j(p_t, \delta_t)$, a "cost index," and other cost shifters:

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

Here η_{jt} is an unobserved cost shock and $\left(z_{jt}^{(1)}, z_{jt}^{(2)}\right)$ are observed cost shifters, with $z_{jt}^{(1)} \in \mathbb{R}$. We permit $z_{jt}^{(2)}$ to include components of $x_{jt}^{(2)}$, although $x_{jt}^{(1)}$ is excluded. We will be explicit below about the independent variation required of $z_{1t}^{(1)}, \ldots, z_{Jt}^{(1)}$.

Parallel to our model of demand, (12) imposes perfect substitution between the unobserved cost shock η_{jt} and the cost shifter $z_{jt}^{(1)}$ inside the nonparametric function mc_j. This is an important restriction, but one that is satisfied in many standard models.[30] We denote the cost index by $\zeta_{jt} \equiv z_{jt}^{(1)} + \eta_{jt}$. As with the parallel assumption on the demand side, the linearity of the index can be relaxed under additional assumptions discussed in Appendix B.

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

Assumption 8. For all j

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

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

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

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

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

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

Given the index restriction, part (i) of Assumption 8 is fairly weak: weak monotonicity in an unobservable could be assumed without loss, since that would merely define an order on the unobservable. Part (ii) requires strictly downward sloping demand. Part (iii) is a high-level condition requiring that it be possible to rewrite first-order conditions to express marginal cost for each product as a function of equilibrium quantities (market shares), prices, and derivatives of demand at these prices and quantities. Our ability to express this matrix in terms of (s_t, p_t) exploits the invertibility result of Lemma 1. We show in Appendix C that, under assumptions already made, part (iii) holds for a variety of standard oligopoly models, including the multi-product price-setting oligopoly model most often used in empirical work. By referring to the high-level condition in part (iii) of Assumption 8 we will be able to provide results for a class of models rather than just one. This will be particularly useful when we discuss discrimination between alternative models.

The following lemma shows that under Assumption 8, firms' first-order conditions imply a unique vector of cost indices ζ_t for any vector of equilibrium prices and market shares.

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

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

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

Given M_t, the right-hand side is an unknown function of s_t, p_t. $\square$

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

(13) as[31]

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

This provides a key set of equations for what follows.

Note that equation (14) takes the same form as (11). We will use this relation in the same way. Let $\tilde{z}_t$ include the exogenous cost shifters $(z_{1t}, \ldots, z_{Jt})$. We can then show the following result.

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

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

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

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

6 A Change of Variables Approach

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

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

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

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

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

6.1 Identification of Demand and the Random Utility Model

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

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

Here we consider these $2J$ equations as a system.

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

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

Although the invertibility results above ensure that there is a unique (δ_t, ζ_t) associated with any (s_t, p_t), the change of variables approach requires that this map be one-to-one. The market share functions (9) 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, ζ_t).

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

This assumption is satisfied if, at the true marginal cost and demand functions, the equilibrium first-order conditions have a unique solution (for prices) given any (δ, ζ). This is often difficult to verify in models of product differentiation (see, for example, Caplin and Nalebuff (1991)), and it is not hard to construct examples in which multiple equilibria do exist. If there are multiple equilibria, Assumption 9 requires an equilibrium selection rule such that the same prices p_t arise whenever (δ_t, ζ_t) is the same. This rules out random equilibrium selection or equilibrium selection based on x_{jt} or ξ_{jt} instead of their sum δ_{jt} (and similarly for ζ_{jt}).

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

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

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

Finally, we add assumptions on the excluded demand and cost shifters. Assumption

12 requires full independence from the structural errors, while Assumption 13 ensures that these instruments have sufficient variation to trace out the demand and cost functions.

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

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

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

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

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

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

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

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

The Jacobian determinants cancel.[32] Thus, fixing $(\hat{s}, \hat{p}, \hat{x}, \hat{z})$, $\phi\left(\hat{s}, \hat{p}, \hat{x}, \hat{z}, x, z\right)$ is equal to the joint density $f_{\xi,\eta}\left(\sigma_1^{-1}\left(\hat{s}, \hat{p}\right) - x_1, \ldots, \pi_J^{-1}\left(\hat{s}, \hat{p}\right) - z_J\right)$ rescaled by the constant denominator

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

in (17). Since this density must integrate (over $(x,z)\in\mathbb{R}^{2J}$) to one, the constant is uniquely determined and the value of

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

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

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

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

This provides a constructive proof of the identification of demand. As with our analysis using general IV conditions, we can extend the identification of demand to obtain full identification of the random utility model under the quasilinear specification of preferences in Assumption 1b. The proof follows that of Theorem 2 and is therefore omitted.

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

6.2 Identification of Marginal Costs

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

Theorem 6. *Suppose that Assumptions 1a, 2, 3, and 8–13 hold. Then, for all* j *(i) each*

η_{jt} is identified and (ii) if each ψ_j is known, the function $mc_j(q_{jt}, \zeta_{jt})$ is identified on the support of $\left(q_{jt}, \zeta_{jt}\right)$.

Proof. Part (i) follows by observing that the argument used in the proof of Theorem 4 can be repeated with trivial modification to recover the inverse pricing relations π_j^{-1} and the cost shocks ζ_{jt}.[33] Part (ii) then follows from the argument used to prove part (ii) of Theorem 3. □

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

7 Discriminating Between Oligopoly Models

A remaining question is whether the correct model of oligopoly competition can be distinguished from alternative models. Bresnahan (1982) offered an influential insight for how "rotations of demand" could be used to do this, citing formal results in Lau (1982) (see also Bresnahan (1989)). While Lau (1982) considered homogeneous goods markets within the context of deterministic conjectural variations models, Bresnahan's original intuition suggested much broader applicability. The following remark shows that a variation of this insight can indeed be extended to our stochastic differentiated products framework.

Remark 1. *Suppose $mc_j(\cdot) = mc_{j'}(\cdot)$ and $\left(q_{jt}, \zeta_{jt}\right) = \left(q_{j't'}, \zeta_{j't'}\right)$, with $t' \neq t$ and/or $j' \neq j$. Under Assumption 8, $\psi_j\left(s_t, D_t\left(s_t, p_t\right), p_t\right) = \psi_{j'}\left(s_{t'}, D_{t'}\left(s_{t'}, p_{t'}\right), p_{t'}\right)$.*

To see this, recall that under Assumption 8

$$mc_j(q_{jt}, \zeta_{jt}) = \psi_j\left(s_t, D_t\left(s_t, p_t\right), p_t\right) \;\; \forall j, t. \tag{19}$$

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

The right-hand side of (19) cannot change unless the left-hand side does. Thus, for example, $(q_{jt}, \zeta_{jt}) = (q_{jt'}, \zeta_{jt'})$ implies $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_j(s_{t'}, D_{t'}(s_{t'}, p_{t'}), p_{t'})$. Likewise, if we assume that $mc_{j'}(\cdot) = mc_j(\cdot)$ for some $j' \neq j$, then $\psi_j(s_t, D_t(s_t, p_t), p_t) = \psi_{j'}(s_t, D_t(s_t, p_t), p_t)$ whenever $(q_{j't}, \zeta_{j't}) = (q_{jt}, \zeta_{jt})$. Thus, Remark 1 provides testable restrictions that can be used to distinguish between alternative models of supply, as long as the conditions for part (i) of Theorem 3 or part (i) of Theorem 6 hold for both models, ensuring identification of the cost shocks ζ_{jt} and $\zeta_{jt'}$.

To illustrate, consider first the simple case of a market with one single-product firm. Consider the null hypothesis that the firm prices at marginal cost and the alternative that the firm is a profit-maximizing monopolist. Figure 2 shows the market demand curve D_t. Under the monopoly hypothesis the function ψ_j in Assumption 8 is the marginal revenue curve MR_t. We label this curve ψ^1_{jt}, 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^0_{jt}. The observed equilibrium outcome E_t in market t maps to two possible values of marginal cost at the quantity q_t, depending on the model.

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

This is a particularly simple example but describes a general "recipe" for ruling out false models using Remark 1. Indeed, the heuristic illustration in Figure 3 applies to any null and alternative oligopoly models. Any false null can be ruled out as long as there exist changes in the market environment that induce rotations of ψ_j for some product j under the true model that are not also rotations under the false null. This is easy to see analytically

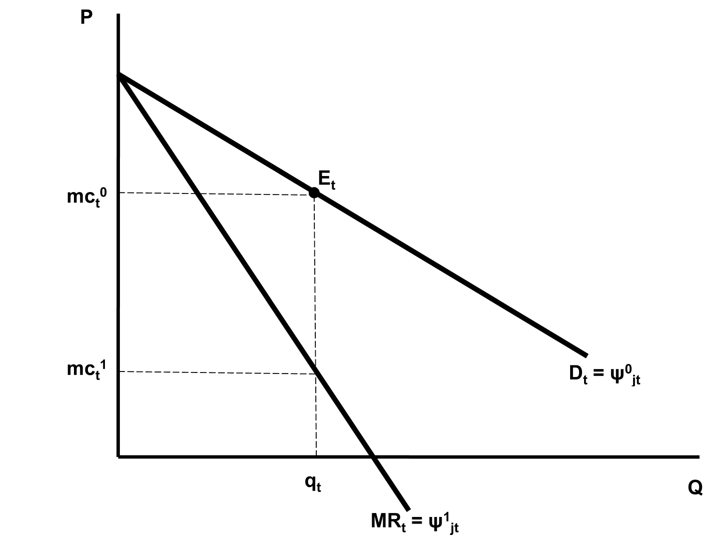

Figure 2: Market outcome E_t maps to different marginal costs under the null and alternative.

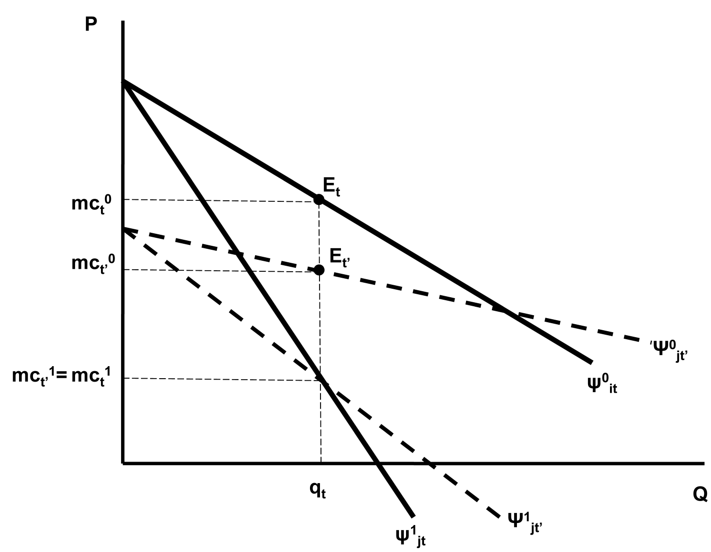

Figure 3: A rotation of the true ψ_j rules out the false null.

as well. Consider two markets t and t' with $\zeta_{jt} = \zeta_{jt'} = \zeta$ and $q_{jt} = q_{jt'} = q$. Since $mc_j\left(q_{jt}, \zeta_{jt}\right) = mc_j\left(q_{jt'}, \zeta_{jt'}\right) = mc_j\left(q, \zeta\right)$, we can use (15) to rationalize $q_{jt} = q_{jt'} = q$ under the null only if ψ_j is the same in both markets under the null, as it is under the true model.

This observation falls directly out of our identification analysis but is very closely related to well known insights from an early literature on identification of firm "conduct" within the class of conjectural variations models (e.g., Bresnahan (1982), Lau (1982)).[34] Our graphical illustration, in particular, is intentionally similar to that given by Bresnahan (1982),[35] but makes explicit the key role of the "residual marginal revenue" function, $\psi_j\left(\cdot\right)$. While we generalize and reinterpret these earlier insights, the message is very similar: one can distinguish between competing models as long as there are changes in the market environment that can shift equilibrium quantity and markup independently, at least for some product. Conditions guaranteeing such variation will depend on the model. However, the changes in the environment that alter ψ_j form a larger set than the changes in aggregate demand considered by Lau (1982) or the "rotations of demand" described by Bresnahan (1982, 1989). For example, even if preferences are identical in markets t and t' (i.e., there is no change in demand), ψ_{jt} and $\psi_{jt'}$ can differ due to variation in the number of competitors, the characteristics of competitors' products ($x^{(1)}_{-jt}$ and/or ξ_{-jt}), or the costs of competing firms ($z^{(1)}_{-jt}$ and/or ζ_{-jt}). And although we have conditioned on $x^{(2)}_t$, $z^{(2)}_t$ and M_t, variation in any of these may also be exploited.

8 Conclusion

We have considered nonparametric identification in a class of differentiated products models used in a growing body of empirical work in IO and other fields of economics. We consid-

[34] In these models products are homogeneous, so the demand side is equivalent to a binary choice model. Our model can be adapted to a homogeneous goods environment by dropping the product (i.e., "j") subscripts on the demand side, interpreting j as a firm subscript on the supply side, and applying the appropriate convention for allocating total output among firms.

[35] Note that once the demand and cost shocks have been identified and held fixed, we have a non-stochastic environment, as in Breshnahan's graphical illustration and the formal results of Lau (1982).

ered two types of arguments.. One links identification of these models to the same kinds of conditions used to show identification of regression models, while the other has close connections to classical identification arguments for supply and demand models. We also pointed to testable implications that can be used to discriminate between alternative models of oligopoly competition.

Our hope is that our results will be useful to both producers and consumers of empirical work on differentiated products markets. The results should help practitioners focus on the essential sources of variation needed to address a wide range of positive and normative questions. For identification of demand or of marginal costs, the critical issue is the availability of instruments. It should be no surprise that there is no getting around the need for instruments, and it should be comforting that this is essentially all that is needed.[36] Likewise our work should help policy makers, managers, and others who might rely on discrete choice demand estimates for making decisions. Our results demonstrate that the nonparametric foundation for empirical work based on these models is really no different from that for simpler, more familiar models. We hope this will aid critical readers in focusing on the key sources of variation in particular applications and ultimately lead to more informed decision making.

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

Appendix A. Connected Substitutes

In this appendix we provide some key lemmas related to the connected substitutes assumption, including the proof of Lemma 1 stated in the text. Here we will repeatedly use the condition "$(\delta, p) \in \chi^{\mathcal{J}}$" as shorthand for "$\left\{x_j, p_j, \xi_j\right\}_{j \in \mathcal{J}} \in \chi^{\mathcal{J}}$ with $\delta_j \equiv x_j + \xi_j \;\; \forall j$."

We begin by stating a useful elementary result in matrix theory (see, e.g., Horn and Johnson (1990), section 6.2).

Lemma 3. *Consider an $n \times n$ matrix A with elements a_{ij}. The following conditions are equivalent:*

(i) the directed graph of A is strongly connected;

(ii) for any strict subset $\mathcal{K} \subset \{1, \ldots, n\}$, there exists $k \in \mathcal{K}$ and $\ell \notin \mathcal{K}$ such that $a_{\ell k} \neq 0$;

(iii) A is irreducible.[37]

The following corollary applies part (ii) of Lemma 3 to our model.

Corollary 1. *Under Assumptions 1a, 2, and 3, for any $(\delta, p) \in \chi^{\mathcal{J}}$ and any strict subset $\mathcal{K} \subset \mathcal{J}$, there exists $k \in \mathcal{K}$ and $\ell \notin \mathcal{K}$ such that $\sigma_\ell(\delta, p)$ is strictly decreasing in δ_k.*

As discussed in the text, the following lemma provides the key implication of the connected substitutes assumption (Assumption 3) for our analysis.

Lemma 4. *Suppose $(\delta, p) \in \chi^{\mathcal{J}}$, $(\delta', p) \in \chi^{\mathcal{J}}$, and $\delta' \neq \delta$. Under Assumptions 1a, 2, and 3, (i) if $I^+ \equiv \left\{j : \delta'_j > \delta_j\right\}$ is nonempty then $\sum_{j \in I^+} \sigma_j(\delta', p) > \sum_{j \in I^+} \sigma_j(\delta, p)$; (ii) if $I^- \equiv \left\{j : \delta'_j < \delta_j\right\}$ is nonempty then $\sum_{j \in I^-} \sigma_j(\delta', p) < \sum_{j \in I^-} \sigma_j(\delta, p)$.*

Proof. Consider part (i) and note that because $0 \notin I^+$, I^+ is a strict subset of $\mathcal{J}$. By Corollary 1, for some $k \in I^+$ and some $\ell \notin I^+$, $\sigma_\ell(\delta, p)$ is strictly decreasing in δ_k. Taking one such pair (k, ℓ), define δ^* by

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

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

By monotonicity of v_{ijt} in δ_{jt} for all j, $\sigma_j(\delta^*, p) \leq \sigma_j(\delta, p)$ for all $j \notin I^+$. Further, because $\ell \notin I^+$

$$\sum_{j \notin I^+} \sigma_j(\delta^*, p) < \sum_{j \notin I^+} \sigma_j(\delta, p).$$

By monotonicity of v_{ijt} in δ_{jt} for all j, $\sigma_j(\delta', p) \leq \sigma_j(\delta^*, p)$ for all $j \notin I^+$, so that we then have

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

Since market shares must sum to one, the result follows. Part (ii) follows from a symmetric argument. □

With these preliminary results, the invertibility of market shares follows easily:

Proof of Lemma 1. Arguing by contradiction, suppose $\delta' \neq \delta$ but $\sigma_j(\delta', p) = \sigma_j(\delta, p)$ for all j. If $\delta'_j \leq \delta_j$ for all $j \neq 0$ then the set $I^- = \left\{j : \delta'_j < \delta_j\right\}$ must be nonempty to satisfy $\delta' \neq \delta$. By Lemma 4 we would then have

$$\sum_{j \in I^-} \sigma_j(\delta', p) < \sum_{j \in I^-} \sigma_j(\delta, p)$$

contradicting the hypothesis that $\sigma_j(\delta', p) = \sigma_j(\delta, p)$ for all j. So it must be that the set $I^+ \equiv \left\{j : \delta'_j > \delta_j\right\}$ is nonempty. But then by Lemma 4 we must have $\sum_{j \in I^+} \sigma_j(\delta', p) > \sum_{j \in I^+} \sigma_j(\delta, p)$, which again contradicts the hypothesis. □

Finally, we state a related result that is used when considering relaxation of the linear index assumption (see Appendix B).

Lemma 5. *Suppose Assumptions 1a, 2, and 3 hold. For all j, $\sigma_j^{-1}(s, p)$ is strictly increasing in s_j.*

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

$$\begin{aligned} s'_1 &> s_1 \\ s'_j &= s_j \quad \forall j > 1 \end{aligned}$$

but

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

Because probabilities sum to one, we must have

$$s_0' \equiv \sigma_0(\delta', p) < \sigma_0(\delta, p) \equiv s_0. \tag{20}$$

If the set $I^+ \equiv \left\{j : \delta_j' > \delta_j\right\}$ 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 $s_j' = s_j \quad \forall j > 1$. Thus, I^+ must be empty, i.e.,

$$\delta_j' \le \delta_j \quad \forall j > 0.$$

But in that case monotonicity of v_{ijt} in δ_j for all j requires

$$\sigma_0(\delta', p) \ge \sigma_0(\delta, p)$$

which contradicts (20). □

Appendix B. Relaxing the Index Restriction

Here we show that, for our results based on nonparametric instrumental variables arguments, the linear index restrictions can be relaxed if we are willing to strengthen the instrumental variables conditions.

For this appendix, suppose $\delta_{jt} = \delta_j\left(x_{jt}^{(1)}, \xi_{jt}\right)$, with δ_j strictly increasing in its second argument. This relaxes the linear separability assumed in the text, but retains the key requirement of monotonicity in ξ_{jt}. As usual, we will condition on a value of $x_t^{(2)}$, suppress these arguments in the notation, and let x_{jt} denote $x_{jt}^{(1)}$.

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

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

for some functions σ_j^{-1}, $j = 1, \ldots, J$. With δ_j strictly increasing in ξ_{jt} we can then write

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

for some function g_j. Moreover, Lemma 5 (see Appendix A) implies that g_j must be strictly increasing in s_{jt}. Thus,

$$\begin{aligned} s_{jt} &= g_j^{-1}\left(\xi_{jt}; s_{-jt}, p_t, x_{jt}\right) \\ &\equiv h_j\left(s_{-jt}, p_t, x_{jt}, \xi_{jt}\right) \end{aligned} \tag{21}$$

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

Now consider the identification of the functions h_j in (21). Because x_{jt} is exogenous, we can condition on it and drop it from the notation, rewriting (21) as

$$s_{jt} = h_j\left(s_{-jt}, p_t, \xi_{jt}\right) \tag{22}$$

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 assume that (s_t, p_t) are continuously distributed conditional on $\tilde{z}_t, x_t$. We then let $f_{jsp}\left(s_{-jt}, p_t | \tilde{z}_t, x_t\right)$ denote the conditional (marginal) density of s_{-jt}, p_t, and let $f_{js}\left(s_{jt} | s_{-jt}, p_t, \tilde{z}_t, x_t\right)$ denote the conditional density of s_{jt}.

Let ϵ_1 and ϵ_2 be some small positive constants. Let

$$\alpha_j(s_{-jt}, p_t) \equiv \{s : f_{js}(s|s_{-jt}, p_t, \tilde{z}_t, x_t) \geq \epsilon_1 \;\forall\, (\tilde{z}_t, x_t) \text{ with } f_{jsp}(s_{jt}, p_t|\tilde{z}_t, x_t) > 0\}.$$

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

Assumption 14. $\xi_{jt} \perp\!\!\!\perp (\tilde{z}_t, x_t)\, \forall t$.

Assumption 15. For all j and $\tau \in (0,1)$,
(i) 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 $m_j(\cdot, \tau) \in \mathcal{L}_j(\tau)$ and $\varepsilon_{jt} \equiv s_{jt} - h_j(s_{-jt}, p_t, \tau)$, $E[B(s_{-jt}, p_t, \tau)\psi(s_{-jt}, p_t, \tilde{z}_t, x_t, \tau)|\tilde{z}_t, x_t] = 0$ a.s. only if $B_j(s_{-jt}, p_t, \tau) = 0$ a.s. for $\psi(s_{-jt}, p_t, \tilde{z}_t, x_t, \tau) = \int_0^1 f_{\varepsilon_j}(bB(s_{-jt}, p_t, \tau)|s_{-jt}, p_t, \tilde{z}_t, x_t)\; db > 0$.
(ii) the conditional density $f_{\varepsilon_j}(e|s_{-jt}, p_t, \tilde{z}_t, x_t)$ 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 14 strengthens the exclusion restriction of Assumption 4 to require full independence instead of mean independence. Assumption 15 is a type of "bounded completeness" condition that replaces Assumption 5 in the text. It was used previously by Chernozhukov and Hansen (2005) to demonstrate nonparametric identification of quantile treatment effects.

Theorem 7. *Suppose Assumptions 1a, 2, 3, 14 and 15 hold. Then for all j, (i) ξ_{jt} is identified for all t, and (ii) the function $s_j(\{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}})$ is identified on $\chi^{\mathcal{J}}$.*

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 (22) is equivalent to the model they consider. Parts (i) and (ii) then follow immediately, as in the proof of Theorem 1. □

This shows that the analog of Theorem 1 (i.e., identification of demand) can be obtained with the relaxed index structure. Extension to full identification of the random utility model under the additional quasilinearity restriction follows exactly as in Theorem 2.

An analogous argument would apply to allow identification of marginal costs under the index structure

$$mc_j\left(q_{jt}, \zeta_{jt}, z_{jt}^{(2)}\right)$$

with

$$\zeta_{jt} = \zeta_j\left(z_{jt}^{(1)}, \eta_{jt}\right)$$

and ζ_j strictly monotone in η_{jt}, relaxing the linear structure $\zeta_{jt} = z_{jt}^{(1)} + \eta_{jt}$ in the text. Because the argument is completely parallel that for identification of demand, we omit it.

Appendix C. Oligopoly First-Order Conditions

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

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

Examples of First-Order Conditions

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

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

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

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

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

which is easily solved for marginal cost:

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

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

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

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

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

$$\sigma_t + \Delta_t (p_t - mc_t) = 0 \tag{23}$$

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

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

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

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

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

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

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

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

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

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

Solutions for Marginal Costs

Price-Setting

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

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

To show invertibility of the matrix Δ_t for the multi-product case, we rely on "Taussky's theorem" which shows that an irreducibly diagonally dominant matrix is invertible.[38] A matrix A is *diagonally dominant* if for all i

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

An *irreducibly diagonally dominant* matrix is a square matrix that is irreducible (see footnote 37) and diagonally dominant, with at least one of the diagonals being strictly dominant, i.e., with at least one row such that

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|. \tag{24}$$

Proposition 1. *For any strict subset of products $\mathcal{K} \subset \mathcal{J}$, re-index the products in $\mathcal{K}$ from 1 to $|\mathcal{K}|$ and let $D(\mathcal{K})$ be the $|\mathcal{K}|$ by $|\mathcal{K}|$ matrix with (k, j) element $\partial \sigma_{kt} / \partial p_{jt}$. Given Assumptions 1a, 2, 3, 7 and parts (i) and (ii) of Assumption 8, $D(\mathcal{K})$ is invertible.*

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

To show this, we begin with the following lemma:

Lemma 6. *Given Assumptions 1a, 2, 3, 7 and parts (i) and (ii) of Assumption 8, for any strict subset $\mathcal{K} \subset \mathcal{J}$, $D(\mathcal{K})$ is a diagonally dominant matrix, with at least one strictly dominant diagonal.*

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

$$\left|\frac{\partial\sigma_{jt}}{\partial p_{jt}}\right| = \sum_{k\in\mathcal{K}-\{j\}} \frac{\partial\sigma_{kt}}{\partial p_{jt}} + \sum_{\ell\notin\mathcal{K}} \frac{\partial\sigma_{\ell t}}{\partial p_{jt}}. \tag{25}$$

By part (ii) of Assumption 8, this implies

$$\left|\frac{\partial\sigma_{jt}}{\partial p_{jt}}\right| \geq \sum_{k\in\mathcal{K}-\{j\}} \frac{\partial\sigma_{kt}}{\partial p_{jt}}. \tag{26}$$

Furthermore, by the connected substitutes assumption, Lemma 3 (see Appendix A), and the strict monotonicity of v_{ijt} in p_{jt}, the second sum in (25) is strictly positive for at least one product $j \in \mathcal{K}$. For that j the inequality in (26) is strict. □

Proof of Proposition 1. We argue that $D(\mathcal{K})$ must be either (i) an irreducibly diagonally dominant matrix, or (ii) block-diagonal with each block being an irreducibly diagonally dominant matrix; then, by Taussky's theorem, $D(\mathcal{K})$ must be invertible. By Lemma 3 (see Appendix A) $D(\mathcal{K})$ is irreducible if and only if the directed graph of $D(\mathcal{K})$ is strongly connected. If the directed graph of $D(\mathcal{K})$ is strongly connected then, by Lemma 6, $D(\mathcal{K})$ is an irreducibly diagonally dominant matrix and is therefore invertible. So now consider the case in which the directed graph of $D(\mathcal{K})$ is not strongly connected. By Corollary 2 and Lemma 8 (both in Appendix D) the directed graph of $D(\mathcal{K})$ can be partitioned into isolated strongly connected subgraphs. The nodes in each isolated subgraph correspond to a subset of products that do not substitute outside of the subset. We can therefore rearrange the order of products, with the products in the first strongly connected subset coming first, the next subset following and so on. The resulting permutation of $D(\mathcal{K})$ is block diagonal, with each block being irreducible by Lemma 3. Further, by Lemma 6, each block is diagonally dominant with at least one strictly dominant diagonal. Therefore, by Taussky's theorem, each block is invertible. This implies that the entire $D(\mathcal{K})$ matrix is invertible. $\square$

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

Quantity setting.

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

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

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

Appendix D

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

Lemma 7. *Suppose Assumptions 1a and 2 hold and that for all $j \in \mathcal{J}$, $\Pr\left(v_{ijt} \leq v | x_{jt}, p_{jt}, \xi_{jt}\right)$ is strictly decreasing and continuous in w_{jt}.*[39] *Then good k substitutes to good ℓ at $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j \in J}$ if and only if good ℓ substitutes to good k at $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j \in J_t}$.*

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

$$\Pr\left(v_{ikt} - \epsilon < v_{i\ell t} < v_{ikt}\right) > 0 \quad \forall \epsilon > 0$$

i.e., letting $d_{i\ell kt} = v_{i\ell t} - v_{ikt}$,

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

Thus, arguing by contradiction, suppose that for some $\epsilon > 0$ and some $\gamma > 0$, we have $\Pr\left(d_{i\ell kt} \in (-\epsilon, 0)\right) > 0$ but $\Pr\left(d_{i\ell kt} \in (0, \gamma)\right) = 0$. For this to hold we must have either (a) $\Pr\left(d_{i\ell kt} > 0\right) = 0$, violating Assumption 2 , or (b) $Pr\left(d_{i\ell kt} > 0\right) > 0$ when supp $d_{i\ell kt}$ excludes $(0, \gamma)$, violating the maintained assumption that the utility differences $(v_{ilt} - v_{ikt})$ have convex support. □

This symmetry of $\Sigma(J_t)$ was not required by our results. However, Lemma 7 demonstrates why in standard models (where v_{ijt} is everywhere continuous in w_{jt}) the directed graph of $\Sigma(J_t)$ is bidirectional (recall Figure 1). Further, the following Corollary is utilized in Appendix C.

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

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

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

Lemma 8. *Let G be a strongly connected directed graph, and let N be a subgraph of G with at least one vertex. If all edges of G are bidirectional, N can be partitioned into isolated strongly connected subgraphs.*

Proof. Let $N_0 = N$, and define $C(n_0)$ to be the null graph (no nodes). Consider an iterative argument, beginning with iteration $t = 1$. Let N_t be the subgraph of N_{t-1} obtained by excluding the nodes of $C(n_{t-1})$. Let n_t be any node of N_t and let $C(n_t)$ be the subgraph of N_t whose nodes are n_t and all nodes from which there is a path in N_t to n_t. By bidirectionality, $C(n_t)$ is strongly connected. If $C(n_t) = N_t$, the argument is complete. Otherwise, add 1 to t and iterate. The argument will be complete in at most $|N|$ iterations. $\square$

References

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

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

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

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

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

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

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

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 60(4), 889–917.

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

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

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

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

BERRY, S. T., AND A. PAKES (2007): "The Pure Characteristics Demand Model," Discussion paper, Yale University.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Gikhman, I. I., and A. V. Skorokhod (1980): *The Theory of Stochastic Processes*. Springer-Verlag, Berlin.

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

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

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

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

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

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

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

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

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

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

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

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

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

Hotz, J., and R. A. Miller (1993): "Conditional Choice Probabilites and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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