Identification of Discrete Choice Demand From Market Level Data*

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

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

March 4, 2009

Abstract

We consider nonparametric identification of random utility discrete choice models of demand for differentiated products. We examine the case of market level data, i.e., observations of product characteristics, market characteristics, and market shares. Our representation of preferences nests random coefficients discrete choice models widely used in practice in the literature on demand for differentiated products; however, our model is nonparametric and distribution free. It allows for choice-specific unobservables, endogenous choice characteristics (e.g., prices), and high-dimensional taste shocks with arbitrary correlation and heteroskedasticity. Using standard conditions from the literatures on mulitnomial choice, nonparametric instrumental variables, and simultaneous equations, we demonstrate the identifiability of demand and of the full random utility model.

*We had helpful early conversations on this topic with Rosa Matzkin and Yuichi Kitamura. We had helpful feedback from participants in the 2008 Game Theory World Congress and the 2008 LAMES.

1 Introduction

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

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 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? This is the question we explore here.

We study the identification of random utility multinomial choice models in the case of "market data," where the researcher observes product characteristics, market characteristics, and market shares, but not individual choices.[1] Our representation of preferences is nonparametric and distribution free, but nests random coefficients discrete choice models widely used

[1] A setting in which one observes individual choices, with or without individual characteristics, is equivalent except when there are observables that shift the relative desirabililty of products in a given choice set. In that case the the additional invididivual level data (i.e., "micro data") can provide additional information, in effect providing variation in the choice set while the choices themselves—in particular, choice-specific unobservables—are held fixed. We address identification in the case of micro data in Berry and Haile (2008).

in practice. It allows for choice-specific unobservables, endogenous choice characteristics, unknown heteroskedasticity, and high-dimensional correlated taste shocks.

We investigate identification of demand as well as identification of the full random utility model. Identification of demand naturally requires instruments for prices (and/or any other endogenous choice characteristics), and we provide one result relying on standard nonparametric instrumental variables conditions (Newey and Powell (2003)). Because identification of demand is sufficient for many purposes motivating estimation of random utility discrete choice models, we view this as an important result. Nonetheless, nonparametric instrumental variables conditions can be difficult to interpret or verify. We provide an alternative result relying instead on a large support condition for the case in which cost-shifters are available as instruments. This result requires additional structure on the supply side, although not a fully specified supply model.

To show full identification of the random utility model we require an additional separability restriction on preferences and a large support condition. Even with this added restriction, our model generalizes those considered previously in the applied and econometrics literatures. And although large support conditions are strong and controversial, they are also standard for evaluating identifiability under ideal conditions. Our results show that full identification of a very general fully nonparametric model can be obtained under the same kind of support conditions previously used to demonstrate identification in even the simplest semiparametric random utility models.

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

In addition to our separate focus on identification of demand, our work is distinguished by the approach of modeling of utility as a nonparametric random function of observed and unobserved choice characteristics. This contrasts with the usual practice of building up randomness from random coefficients and/or other taste shocks. Our approach facilitates focusing directly on the identifiability of the joint distribution of utilities (conditional on observed and unobserved characteristics) rather than on the joint distribution of taste shocks. This is a significant advantage because the vector of utilities has the same dimension as the vector of observable choice probabilities; thus, the dimension of the model primitives equals that of the observables. This simple fact enables us to obtain positive identification results for a very general class of models.

Essential to our model is an unobservable associated with each choice and market. Although explicit modeling of these choice-specific unobservables is standard in the applied literature, most prior econometric work on discrete choice with endogeneity has embedded the sources of randomness in preferences and the sources of endogeneity in the same random variables. Explicitly modeling choice-specific unobservables can be important for defining the objects of interest, particularly the effects of changes in endogenous characteristics within a flexible model of consumer heterogeneity. For example, our formulation allows us to characterize demand elasticities, which require evaluating the effects of a change in price (including resulting changes in the variance or other moments of random utilities) holding unobserved product characteristics fixed.[2]

In the following section we briefly place our work in the context of the prior literature. We set up the model in section 3, then discuss a key preliminary result in section 4. We provide our two sets of identification results in sections 5 and 6. We conclude in section 7.

2 Related Literature

In addition to the large applied literature that motivates our work, the ideas in this paper relate to several literatures. Attention to the identifiability of discrete choice models has a long history. Important early work includes Manski (1985), Manski (1988), Matzkin

[2] See, e.g., Chintagunta (2001) for evidence of the importance of allowing for both heterogeneity and endogeneity.

(1992), and Matzkin (1993), who examined semiparametric models with exogenous regressors. Work considering identification of heterogeneous preferences has focused on random coefficients models and includes Ichimura and Thompson (1998) and Gautier and Kitamura (2007), which focus on binary choice. Briesch, Chintagunta, and Matzkin (2005) consider multinomial choice, allowing generalizations of the linear random coefficients model. Our work relaxes functional form and distributional assumptions in this earlier work, but we exploit many of the same insights, particularly the reliance on exogenous observables to trace out the distribution of unobservables.

Identification results for discrete choice models allowing for endogeneity have been given by Honoré and Lewbel (2002), Hong and Tamer (2004), Lewbel (2000), Lewbel (2005), Blundell and Powell (2004), and Magnac and Maurin (2007). These all consider linear semiparametric models in which an additive scalar shock (analogous to the extreme value or normal shock in logit and probit models) may be correlated with some observables. Among these, Lewbel (2000) and Lewbel (2005) consider multinomial choice. Extensions to nonadditive shocks are considered in Matzkin (2007b) and Matzkin (2007a).

In addition to relaxing functional form restrictions on the roles of observables and unobservables, relative to this prior literature we add a distinction between choice-specific unobservables and individual heterogeneity in preferences.[3] As discussed in the introduction, this is important in a number of applications if one is to construct counterfactual predictions (e.g., responses to changes in price) that account for both heteroskedasticity (e.g., heterogeneity in the marginal rate of substitution between income and other characteristics) and endogeneity (e.g., correlation between a good's price and its unobserved quality). This separate treatment of choice-specific unobservables and preference heterogeneity builds on parametric models used in applied work, including that in Berry, Levinsohn, and Pakes (1995) and a large related literature. Matzkin (2004) (section 5.1) also makes a distinction between choice-specific

[3]Concurrent work by Fox and Gandhi (2008) explores identifiability of several related models, including a flexible model of multinomial choice in which consumer types are multinomial and utility functions are determined by a finite parameter vector. They suggest that one of our approaches for allowing choice-specific unobservables and endogenous choice characteristics could be adapted to their framework.

unobservables and an additive preference shock, but without heteroskedasticity.[4]

Blundell and Powell (2004), Matzkin (2004), and Hoderlein (2008) have considered binary choice in semiparametric triangular models, leading to the applicability of control function methods or the related idea of "unobserved instruments."[5] For binary choice demand estimation, triangular models can be appropriate when price depends only on the unobserved demand shock, but not on a cost shock as well. In the case of multinomial choice, standard models of oligopoly pricing in differentiated products markets imply that each price depends on the entire vector of demand shocks (and cost shocks, if any). This appears to preclude the use of control function methods; however, one of our results below uses a related approach of inverting a multivariate supply and demand system to recover the entire vector of shocks to supply and demand.

Our attention to models in which prices and market shares are determined simultaneously naturally leads us in some cases to ideas explored previously for identification of simultaneous equations models (e.g., Brown (1983), Roehrig (1988), Benkard and Berry (2006), Matzkin (2005), and Matzkin (2008)). Matzkin (2008) in particular has recently explored identification in a variety of nonparametric simultaneous equations models. Although she does not explicitly address discrete choice models, for one of our results we transform the model to a form equivalent to one she considers. Our assumptions and proof for this case differ from hers in important ways.[6] Nonetheless, we exploit a separability condition whose advantages she also emphasizes. Our transformation of the discrete choice model to the simultaneous equations setup applies the insights of Gandhi (2008), who recently showed how to extend a key invertibility result of Berry (1994) and Berry and Pakes (2007) to a more general class of models.

Finally, our own work in Berry and Haile (2008) explores identification in the case of "micro data" using ideas similar to some of those we use. The distinction between "market

[4]See also Matzkin (2007a) and Matzkin (2007b).

[5]See also Petrin and Train (2009) and Altonji and Matzkin (2005).

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

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 (2008) is that within a market the choice-specific unobservables are held fixed; therefore, examining how choice probabilities change within a market as consumer-choice-specific observables vary can provide a great deal of information about the randomness of utility, holding fixed the choice-specific unobservable.[7] That strategy was exploited throughout Berry and Haile (2008), but cannot be applied in the case of market data.[8]

3 Model

3.1 Random Utility Discrete Choice

Each consumer i in market t chooses from a set $\mathcal{J}_t$ of available products. The term "market" here is synonymous with the choice set. In particular, consumers facing the same choice set can be viewed as being in the same market. In practice, markets will typically be defined geographically and/or temporally. The choice set $\mathcal{J}_t$ 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 $J_t = |\mathcal{J}_t| - 1$.[9]

Each inside good 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 only because we will allow p_{jt} to be endogenous but will assume other product characteristics are exogenous. The restriction to a single endogenous characteristic reflects the usual practice, but is not essential for our results. Indeed, the

[7] Because the prior econometrics literature has given little attention to choice-specific unobservables, the importance of the distinction between micro data and market data may not have been fully recognized.

[8] Berry and Haile (2008) includes a discussion of a particular case in which what appears to be a "market data" environment is actually isomorphic to the "micro data" environment. The key is that in that example one has a continuum of observations for which the choice-specific unobservables are held fixed while observables vary. In general this is not the case.

[9] 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.

results in sections 4 and 5 hold without modification whatever the dimension of p_{jt}.[10] We allow x_{jt} to include product dummies; thus, utilities may depend on the "names" of the products as well as their characteristics. Unobserved product/market characteristics are represented by an index $\xi_{jt} \in \mathbb{R}$. A market t is thus characterized by $\left(\mathcal{J}_t, \left\{x_{jt}, p_{jt}, \xi_{jt}\right\}_{j \in \mathcal{J}_t}\right)$. We let $\chi = \mathrm{supp}\left(x_{jt}, p_{jt}, \xi_{jt}\right)$.

We consider preferences represented by a random utility model. Each consumer i in market t has a conditional indirect utility function $u_{it} : \chi \to \mathbb{R}$. However, consumers have heterogeneous tastes, even conditional on all observables. From our perspective, each consumer's utility function u_{it} is a random draw from a set $\mathcal{U}$ of permissible functions $\{\tilde{u} : \chi \to \mathbb{R}\}$. We discuss restrictions on the set $\mathcal{U}$ below.

Formally, we define the random function u_{it} as follows. 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 $\omega \in \Omega$.

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, we emphasize that ω_{it} is not a random variable but an elementary event that can determine an arbitrary number of random variables. The following example illustrates.

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

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

where, for example, the random variables $(\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ *are defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *as* $\left(\alpha\left(\omega_{it}\right), \beta^{(1)}\left(\omega_{it}\right), \ldots, \beta^{(K_x)}\left(\omega_{it}\right), \epsilon_1\left(\omega_{it}\right), \ldots, \epsilon_J\left(\omega_{it}\right)\right)$.

[10] The modifications required for the results in section 6 are straightforward.

As is standard, without loss we assume the draw ω_{it} from the sample space Ω determining the function u_{it} is independent of the arguments of the function, $\left(x_{jt}, p_{jt}, \xi_{jt}\right)$. We add to this the standard assumption of menu-independent preferences.[11]

Assumption 1. The measure $\mathbb{P}$ on Ω does not vary with $\mathcal{J}_t$ or $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j\in\mathcal{J}_t}$.

Relative to the standard formulation of a random function on χ, this merely rules out the possibility that the utility from one product varies with the set or characteristics of other products on offer. This is a standard assumption of stable preferences made in nearly all work on discrete choice.[12]

Aside from this menu-independence and the restriction to scalar choice-specific unobservables, our representation of preferences is so far fully general. For example, it allows arbitrary correlation of consumer-specific tastes for different goods or characteristics. It also allows arbitrary heteroskedasticity in utilities across different elements of $\mathcal{J}_t$, or in utilities for the same element as $\left(x_{jt}, \xi_{jt}\right)$ varies. As the following example shows, even with additional assumptions that we will not make, this structure provides a significant generalization of models typically considered in the literature.

Example 1 (continued). *With the specification*

$$(\alpha_{it}, \beta_{it}, \epsilon_{i1t}, \ldots, \epsilon_{iJt}) = \left(\alpha\left(\omega_{it}\right), \beta^{(1)}\left(\omega_{it}\right), \ldots, \beta^{(K_x)}\left(\omega_{it}\right), \epsilon_1\left(\omega_{it}\right), \ldots, \epsilon_J\left(\omega_{it}\right)\right)$$

Assumption 1 allows an arbitrary joint distribution of $\left(\alpha_{it}, \beta_{it}^{(1)}, \ldots, \beta_{it}^{(K_x)}, \epsilon_{i1t}, \ldots, \epsilon_{iJ_tt}\right)$ *but requires it to be the same for all* i, t, *and* $\left\{\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right\}_{j=1\ldots J}$. *Note that more general specifications are permissible, even within the linear random coefficients model. For ex-*

[11] 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. Note that with Assumption 1 the joint distribution of $\{\epsilon_{ijt}\}_{j\in\mathcal{K}}$ will be the same regardless of whether $\mathcal{K} = \mathcal{J}_t$ or $\mathcal{K} \subset \mathcal{J}_t$. Thus, for example, a consumer's preference between two products j and k does not depend on the other products in the the choice set.

[12] See Block and Marschak (1960), Falmagne (1978), and Barbera and Pattanaik (1986) for discussion of testable implications in the context of discrete choice, and Haile, Hortaçsu, and Kosenok (2008) for an extension to strategic environments.

ample, we can allow the specification $\epsilon_{ijt} = \epsilon_{ijt}(x_{jt}, \omega_{it})$, where Assumption 1 requires that the joint distribution of $(\epsilon_{i1t}, \ldots, \epsilon_{iJ_t t})$ be the same for choice sets with identical observable characteristics (which could include product dummies).

We will now restrict the set of utility functions we permit with a restriction on the set $\mathcal{U}$. First partition x_{jt} into $\left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$ with $x_{jt}^{(1)} \in \mathbb{R}$. We then make the following assumption.

Assumption 2. $\mathcal{U}$ is the set of all functions $\tilde{u}_{it} : \chi \to \mathbb{R}$ such that conditional on any $x_{jt}^{(2)}$, $\tilde{u}_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right) = \mu_{it}\left(x_{jt}^{(1)} + \xi_{jt}, x_{it}^{(2)}, p_{jt}\right)$ for some function μ_{it} that is strictly increasing in its first argument.

With Assumption 2, we consider random utility functions with representations of the form

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

Assumption 2 contains two parts. First is a standard restriction to a vertical product-specific unobservable ξ_{jt}; thus all consumers agree that an increase in ξ_{jt} makes choice j more attractive, all else equal. The second is perfect substitutability between ξ_{jt} and $x_{jt}^{(1)}$. In the the standard linear random coefficients model (Example 1) this holds if there is one covariate that enters without a random coefficient. To foreshadow the role that these two assumptions play below, note that identification requires recovering each latent ξ_{jt}. Monotonicity in ξ_{jt} will provide an essential invertibility condition, while perfect substitutability provides a means of quantifying variation in ξ_{jt} in units of the observable $x_{jt}^{(1)}$.

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

$$u\left(x_{jt}^{(1)} + \xi_{jt}, x_{it}^{(2)}, p_{jt}, \omega_{it}\right) > u\left(x_{kt}^{(1)} + \xi_{kt}, x_{it}^{(2)}, p_{kt}, \omega_{it}\right) \quad \forall k \in \mathcal{J}_t - \{j\}. \tag{3}$$

We denote consumer i's choice by

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

This leads to market shares (choice probabilities)

$$
\begin{aligned}
s_{jt} &= E_{\mathbb{P}}\left[1\left\{y_{it}=j\right\}\right] \\
&= E_{\mathbb{P}}\left[1\left\{u\left(x_{jt}, p_{jt}, \xi_{jt}, \omega_{it}\right) > u\left(x_{kt}, p_{kt}, \xi_{kt}, \omega_{it}\right)\ \forall k \in \mathcal{J}_t - \{j\}\right\}\right] \\
&= s_j\left(\mathcal{J}_t, \left\{x_{kt}, p_{kt}, \xi_{kt}\right\}_{k \in \mathcal{J}_t}\right) \qquad (4)
\end{aligned}
$$

3.2 Normalizations

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 utilities, which have no natural location or units (scale). Throughout the paper we will normalize the location of utilities by setting the utility from the outside good to zero:

$$v_{i0t} = 0.$$

We will be able to define our scale normalization on utility below.

The second type of normalization is a normalization of the choice-specific unobservables ξ_{jt}. The linear substitutability between $x_{jt}^{(1)}$ and ξ_{jt} required by Assumption 2 already defines the scale of each ξ_{jt}. Our choice of location normalization for ξ_{jt} will be defined below.

3.3 Observables and Primitives of Interest

Identification (of demand or the random utility model) with endogenous prices will require excluded instruments, which we denote by $\tilde{z}_{jt}$. Below we will discuss the types of instruments we have in mind as well as the formal properties defining valid instruments. The set of observables then consists of $(t, \mathcal{J}_t, \{s_{jt}, x_{jt}, p_{jt}, \tilde{z}_{jt}\}_{j \in \mathcal{J}_t})$. As usual, to discuss identification, we treat their population joint distribution as known.

We consider two types of identification results. One is *identification of demand*; i.e., of the functions s_j defined in (4). These functions 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.

We also consider identification of the joint distribution of indirect utilities conditional on the choice set $(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$. These conditional distributions are the primitives of the model of consumer choice. We refer to this as *full identification of the random utility model*. Interest in full identification of random utility model rather than merely demand derives from welfare questions. However, the specification of preferences in (2) does not permit meaningful aggregate welfare measures to be defined. This is one reason we will consider full identification of the random utility model under a more restrictive specification of the class of utility functions:

Assumption 2′. $\mathcal{U}$ is the set of all functions $\tilde{u}_{it} : \chi \to \mathbb{R}$ such that for each $x_{jt}^{(2)}$ there is a monotonic transformation Γ such that $\Gamma\left(\tilde{u}_{it}\left(x_{jt}, p_{jt}, \xi_{jt}\right)\right) = x_{jt}^{(1)} + \xi_{jt} + \mu_{it}\left(x_{it}^{(2)}, p_{jt}\right)$ for some function μ_{it}.

This alternate assumption leads to specifications of random conditional indirect utilities with representations of the form

$$v_{ijt} = x_{jt}^{(1)} + \xi_{jt} + \mu\left(x_{jt}^{(2)}, p_{jt}, \omega_{it}\right) \tag{5}$$

which is a special case of (2). This specification incorporates quasilinearity, enabling characterization of utilitarian social welfare.[13]

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

$$v_{ijt} = u_j\left(x_{jt}^{(1)} + \xi_{jt}, p_{jt}, \omega_{it}\right) \tag{6}$$

[13]Since price enters the function μ_j the quasilinearity will not be in wealth. This nonetheless provides a valid notation of utilitarian social welfare in terms of a numeraire characteristic.

and

$$v_{ijt} = x_{jt}^{(1)} + \xi_{jt} + \mu_j \left(p_{jt}, \omega_{it} \right) \tag{7}$$

to represent, respectively, (2) and (5) above.

4 Preliminaries: Inverting Market Shares

Let

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

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

The key implication of Assumption 2 is that choice probabilities depend on the indices δ_{jt}, rather than separately on the components x_{jt} and ξ_{jt}. In particular, for any vector δ_t, market shares are given by

$$\begin{aligned} s_{jt} &= E_{\mathbb{P}} \left[1 \left\{ u_j \left(\delta_{jt}, p_{jt}, \omega_{it} \right) > u_k \left(\delta_{kt}, p_{jt}, \omega_{it} \right) \ \forall k \in \mathcal{J}_t - \{j\} \right\} \right] \\ &= \int_{\mathbf{v}: v_{ijt} \geq v_{ikt} \forall k} dF_v \left(v | \delta_t, p_t \right) \\ &\equiv \sigma_j \left(\delta_t, p_t \right) \end{aligned} \tag{8}$$

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

Following Gandhi (2008), we will assume the follow "strong substitutes" condition.

Assumption 3. Consider any δ such that $\sigma_j \left(\delta, p \right) > 0$ for all $j \in \mathcal{J}$. For any strict subset $\mathcal{K} \subset \mathcal{J}$, there exists $k \in \mathcal{K}$ and $j \notin \mathcal{K}$ such that $\sigma_j \left(\delta, p \right)$ is strictly decreasing in δ_k.

Given the monotonicity of v_{ijt} in δ_{jt}, this is a natural regularity condition requiring that for every binary partition of $\mathcal{J}$ there is some substitution between cells of the partition. This is guaranteed if there are always consumers on the margin of indifference between every pair of choices, as in multinomial probit and logit models (with or without random coefficients). However, that degree of substitution is stronger than we require. For example, in a pure vertical model (e.g., Shaked and Sutton (1982)) each product substitutes with at most two others, yet Assumption 3 holds.[14] With this assumption we can follow the argument in Theorem 2 of Gandhi (2008) to show the following lemma, which generalizes the well-known invertibility results for linear discrete choice models from Berry (1994) and Berry and Pakes (2007).[15]

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

Proof. Suppose, contrary to the claim, that for some $\delta \neq \delta'$, $\sigma_j(\delta, p) = \sigma_j(\delta', p) = s_j$ for all j. Since we have normalized the utility of the outside good to zero for all choice sets, we can define $\delta_0 = \delta_0^{'} = 0$ as a notational convention without loss. Without loss, let $\delta_j' > \delta_j$ for choice $j \in \mathcal{J}$. Because $0 \in \mathcal{J}$, there must then exist a strict subset of choices $\mathcal{K} \subset \mathcal{J}$ such that $\delta_j' > \delta_j \forall j \in \mathcal{K}$ and $\delta_j' \leq \delta_j \forall j \in \mathcal{J} - \mathcal{K}$. For this subset $\mathcal{K}$ let $k \in \mathcal{K}$ be the index of a product referred to as "k" in Assumption 3. Now define a new vector δ^* by

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

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

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

[14] For example, consider a 5 good vertical model and $\mathcal{K} = \{3, 4, 5\}$. Good 4 does not substitute outside of $\mathcal{K}$, but goods 3 and 4 do. Thus the condition holds. It requires only that there be some element of $\mathcal{K}$ that substitutes with some good outside $\mathcal{K}$. This will be true here for any proper subset $\mathcal{K} \subset \mathcal{J} = \{0, 1, \ldots, 5\}$.

[15] 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 since, 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 gauranteeing 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.

so that (since probabilities must sum to one)

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

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

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

which contradicts the hypothesis $\sigma_j\left(\delta,p\right)=\sigma_j\left(\delta',p\right)=s_j$ for all j. □

5 Identification with General IV Conditions

Recall that $\delta_{jt}\equiv x_{jt}+\xi_{jt}$ and that market shares are given by $s_{jt}=\sigma_j\left(\delta_t,p_t\right)$. Using the inversion result of Lemma 1, we have

$$\delta_{jt}=\sigma_j^{-1}\left(s_t,p_t\right)\quad\forall j \tag{9}$$

which we can rewrite as

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

Let $z_{jt}=(\tilde{z}_{jt},x_{1t},\ldots,x_{Jt})$ denote the exogenous variables, where $\tilde{z}_{jt}$ represents instruments for p_{jt} excluded from the determinants of v_{ijt}. Possible instruments include cost shifters, exogenous characteristics of substitute goods (Berry, Levinsohn, and Pakes (1995)), and prices of the same good in other markets (Hausman (1996)). We let z_t denote the matrix $(z_{1t},\ldots,z_{Jt})$.

We first consider identification of demand under a pair of instrumental variables assumptions, which we take from Newey and Powell (2003).

Assumption 4. $E[\xi_{jt}|z_{jt}]=E[\xi_{jt}]$.

Assumption 5. For any function $B(s_t, p_t)$ with finite expectation, $E[B(s_t, p_t)|z_{jt}] = 0$ almost everywhere implies $B(s_t, p_t) = 0$ almost everywhere.

The first is a standard exclusion restriction, requiring mean independence between the instruments and the structural error ξ_{jt}. The second is a "bounded completeness" condition, which is the nonparametric analog of the standard rank condition for linear models, here extended to nonparametric models with separable errors. Lehman and Romano (2005) gives standard sufficient conditions, and additional discussion of this condition can be found in Newey and Powell (2003) and Severini and Tripathi (2006). Roughly speaking, 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.

Theorem 1. *Under Assumptions 1, 2, and 3–5, the functions* $s_j(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$ *are identified at all points of support.*

Proof. For any j, rewriting (10) and taking expectations conditional on z_{jt}, we obtain

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

so that by Assumption 4,

$$E\left[\sigma_j^{-1}(s_t, p_t)\middle| z_{jt}\right] - x_{jt} = \kappa$$

almost everywhere for some constant κ. Suppose there is another function $\tilde{\sigma}_j^{-1}$ satisfying

$$E\left[\tilde{\sigma}_j^{-1}(s_t, p_t)\middle| z_{jt}\right] - x_{jt} = \kappa$$

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

$$E\left[B(s_t, p_t)|z_{jt}\right] = 0 \quad \text{a.e.}$$

But by Assumption 5 this requires $\tilde{\sigma}_j^{-1} = \sigma_j^{-1}$ almost everywhere, implying that σ_j^{-1} is identified. Each ξ_{jt} is then uniquely determined by (10). Since choice probabilities are observed, and all arguments of the demand functions $s_j(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$ are now known, the result follows. $\square$

To obtain full identification, we will require a large support condition on x_t.

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

This support condition is strong. However, 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.[16] This provides a natural benchmark for evaluating identification under ideal conditions on observables, and the following result shows that the addition of this condition suffices to obtain full identification under the representation of preferences in (5).

Theorem 2. *Under Assumptions 1, 2′, and 3–6, the joint distribution of* $(v_{i1t}, \ldots, v_{iJt})$ *conditional on any* $(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k\in\mathcal{J}_t})$ *in their support is identified.*

Proof. With Assumption 2′ the market share of the outside good, conditional on p_t, x_t, and $(\xi_{1t}, \ldots, \xi_{Jt})$ is

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

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

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

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

[16]To our knowledge, all results showing semiparametric or nonparametric identifiation 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 (2008)).

Theorems 1 and 2 provide sufficient conditions for identification of demand and of the full random utility model. The conditions for the latter include quasilinearity of utility, so that meaningful aggregate welfare measures can be defined.

A limitation of these results is their reliance on a high-level instrumental variables condition (Assumption 5), which may be difficult to verify in particular models. Newey and Powell (2003) show that this condition is, in general, necessary for identification of nonparametric regression models, so there is little hope that weaker conditions could be found without imposing additional structure. However, as we show in the following section, we can provide an alternative result when we have cost shifters and are willing to place some weak structure on the supply side. Moreover, in this case we can provide a constructive identification argument.

6 Cost Shifters: A Change of Variables Approach

Here we combine the random utility discrete choice model of demand with some restrictions on supply. This enables us to proceed without Assumptions 4 and 5, instead relying on a constructive argument using a change of variables technique often useful in simultaneous equations models (e.g., Matzkin (2005), Matzkin (2008)). However, to employ the change of variable technique, we will have to strengthen the statistical assumption on the relationship between the exogenous variables and the market unobservables. In particular, we will replace mean independence, assumption 4, with full independence.

Before turning to further assumptions, we discuss the supply-side conditions that are necessary for the change of variables approach. As noted, we make a high-level assumption on the invertibility of the first order conditions for optimal prices. This is a generalization of the "supply-side" inversion of Berry, Levinsohn, and Pakes (1995), which uses first-order conditions for prices to solve for shocks to marginal cost.

6.1 Supply

For simplicity consider the case of single-product firms and let firm j produce product j. We assume firm j's costs have the form

$$C_j\left(q_{jt}, z_{jt}^{(1)}+\eta_{jt}, z_{jt}^{(2)}\right) \tag{11}$$

where q_{jt} is the quantity sold by firm j in market t, η_{jt} is an unobserved cost shock, $\left(z_{jt}^{(1)}, z_{jt}^{(2)}\right)$ are cost shifters, and $z_{jt}^{(1)} \in \mathbb{R}$. We permit $z_{jt}^{(2)}$ to include components of $x_{jt}^{(2)}$; we will be explicit below about our assumptions on the independent variation required of $z_{1t}^{(1)}, \ldots, z_{Jt}^{(1)}$. The specification (11) is our main restriction on the supply side. It requires perfect substitution between the unobserved cost shock and an observable cost shifter. This is a nontrivial restriction analogous to that we made previously on utilities. However, the specification of costs is otherwise unrestricted.

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

If M_t is the measure of consumers in market t, $q_{jt} = M_t\sigma_j\left(\delta_t, p_t\right)$, so we will write $q_{jt}\left(\delta_t, p_t\right)$. Let

$$\zeta_{jt} = z_{jt} + \eta_{jt}.$$

Firm j's profit is then given by

$$p_{jt}q_{jt}\left(\delta_t, p_t\right) - C_j\left(q_{jt}\left(\delta_t, p_t\right), \zeta_{jt}\right) \tag{12}$$

Rather than assuming a particular extensive form for competition on the supply side, we make the following high-level assumption on equilibrium prices.

Assumption 7. Given any $(\delta_{1t}, \ldots, \delta_{Jt})$, there is at most one vector $(\zeta_{1t}, \ldots, \zeta_{Jt})$ consistent with any given vector of prices $(p_{1t}, \ldots, p_{Jt})$.

There are two parts to this assumption. One is that perfect substitutability between

x_{jt} and ξ_{jt} (and between z_{jt} and η_{jt}) is preserved by the equilibrium supply correspondence. This only rules out equilibrium selection based on x_{jt} or ξ_{jt} instead of their sum δ_{jt} (similarly for ζ_{jt}). With this we can write

$$p_{jt} \in \rho_j \left(\delta_{1t}, \ldots, \delta_{Jt}, \zeta_{1t}, \ldots, \zeta_{Jt}\right) \quad \forall j$$

which (given the perfect substitutability) merely represents prices with completely general reduced form correspondences with the exogenous variables.

The second part of Assumption 7 is uniqueness of the vector $(\zeta_{1t}, \ldots, \zeta_{Jt})$ that rationalizes a given vector of prices, conditional on $(\delta_{1t}, \ldots, \delta_{Jt})$. This is satisfied in standard differentiated products oligopoly models. For example, as the Proposition below shows, Assumption 7 holds in the cases of price or quantity setting under natural assumptions of downward sloping demand and nontrivial marginal cost shifters. The idea is simple: if firm j's first-order condition holds at one value of ζ_{jt}, an increase in ζ_{jt} changes its marginal cost; this will cause its first-order condition to be violated if all firms' behavior is held fixed. Thus, conditional on $(\delta_{1t}, \ldots, \delta_{Jt})$, a given price vector cannot be generated by equilibrium behavior at two different values of the vector $\zeta_t = (\zeta_{1t}, \ldots, \zeta_{Jt})$.

Proposition 1. *Suppose (i) each function σ_j is strictly decreasing and differentiable in p_{jt}, (ii) each function C_j is differentiable in q_{jt}; and (iii) $\frac{\partial C_j(q_{jt},\zeta_{jt})}{\partial q_{jt}}$ is strictly monotonic in ζ_{jt}. Then in a complete information game of price setting or quantity setting, Assumption 7 holds.*

Proof. First consider the price setting game. Suppose to the contrary that there exist two different vectors ζ' and ζ'' leading to the same vector of equilibrium prices $(p_1, \ldots, p_J)$. Without loss, let $\zeta'_j \neq \zeta''_j$. Differentiating (12) with respect to p_j gives the first-order condition

$$q_j(\delta, p) + \frac{\partial q_j(\delta, p)}{\partial p_j}\left[p_j - \frac{\partial C_j\left(q_j, \zeta_j\right)}{\partial q_j}\right] = 0. \qquad (13)$$

This condition must hold at both values of ζ_j, with all else in (13) held fixed. Since demand is strictly downward sloping, this is possible only if the marginal cost $\frac{\partial C_j(q_j,\zeta_j)}{\partial q_j}$ is the same at both values of ζ_j, contradicting hypothesis (iii) of the proposition.

Now consider the quantity setting game. Prices adjust to clear the market. Letting $q = (q_1, \ldots, q_J)$ we can write $p_j = \rho_j(q, \delta)$. Firm j's profit is

$$\rho_j(q, \delta) q_j - C_j(q_j, \zeta_j)$$

Again arguing by contradiction, suppose that the first-order condition

$$\rho_j(q, \delta) + \frac{\partial \rho_j(q, \delta)}{\partial q_j} q_j - \frac{\partial C_j(q_j, \zeta_j)}{\partial q_j} = 0$$

holds at two distinct values of ζ_j. Because demand slopes down, this is possible only if the marginal cost $\frac{\partial C_j(q_j,\zeta_j)}{\partial q_j}$ is the same at both values of ζ_j, contradicting hypothesis (iii) of the proposition. □

We have seen that standard noncooperative oligopoly models imply Assumption 7. Our results will not require knowledge of the true model of competition—only that Assumption 7 holds. This is an advantage of relying on the high-level Assumption 7.[17]

The key implication of Assumption 7 is invertibility, analogous to that guaranteed on the demand side by Lemma 1. In particular, for any (p_t, δ_t) we may write

$$z_{jt} + \eta_{jt} = \rho_j^{-1}(p_{1t}, \ldots, p_{Jt}, \delta_{1t}, \ldots, \delta_{Jt}).$$

Substituting from (9), we have

$$z_{jt} + \eta_{jt} = \rho_j^{-1}\left(p_{1t}, \ldots, p_{Jt}, \sigma_1^{-1}(s_t, p_t), \ldots, \sigma_J^{-1}(s_t, p_t)\right) \quad \forall j.$$

[17] Although our focus is on identification of demand, full treatment of the identifiability of the marginal cost function would require committing to further assumptions on the appropriate model of supply – for example, assuming Nash equilibrium in prices as in Berry, Levinsohn, and Pakes (1995). With such assumptions, showing identification would be straightforward, since the results below already deliver identification of cost shocks.

which we rewrite as[18]

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

6.2 Identification

From the analysis above we take the two equations (10) and (14), which we repeat here:

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

Note that the linear structure normalizes the scale of the unobservables ξ_{jt} and η_{jt} already. To normalize locations, 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

$$\sigma_j^{-1}\left(s^0, p^0\right) - x_j^0 = \pi_j^{-1}\left(s^0, p^0\right) - z_j^0 = 0. \tag{15}$$

We now assume existence of a joint density for the structural errors $(\xi_{1t}, \ldots, \xi_{Jt}, \eta_{1t}, \ldots, \eta_{Jt})$ and a large support for (x_t, z_t).

Assumption 8. The distribution of $(\xi_1, \ldots, \xi_J, \eta_1, \ldots, \eta_J)$ is absolutely continuous with respect to the Lebesgue measure.

Assumption 9. $\text{supp}(x_t, z_t) = \mathbb{R}^{2J}$.

Finally, our change of variables approach requires that (x_t, z_t) be fully independent of the error (ξ_t, η_t)

Assumption 10. $(x_t, z_t) \perp\!\!\!\perp (\xi_t, \eta_t)$.

Theorem 3. *Suppose Assumptions 1, 2, 3, and 7–10 hold. Then the functions* $s_j\left(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}_t}\right)$ *are identified at all points of support.*

[18] This notation is somewhat strained, as we have not defined any function π_j for which π_j^{-1} is the inverse. We write it this way nonetheless as a reminder that this is the result of inverting the equilibrium pricing relation.

Proof. We observe the joint density of the market shares and prices, conditional on the vectors x_t and z_t. By standard change of variables, 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 the matrix $\mathbb{J}\left(s_t, p_t\right)$ is the Jacobian of the vector function $\left(\sigma_1^{-1}, \ldots, \sigma_J^{-1}, \pi_1^{-1}, \ldots, \pi_J^{-1}\right)'$ at the point (s_t, p_t). Note that here we have used the assumption (Assumption 10)that the distribution of (ξ_t, η_t) is the same for all x_t, z_t. Fixing market shares and prices at (s^0, p^0) and letting (x_t, z_t) vary, we observe the values of $f_{s,p}\left(s^0, p^0 | x_t, z_t\right)$, from which we learn the ratios

$$\frac{f_{\xi,\eta}\left(\sigma_1^{-1}\left(s^0, p^0\right) - x_{1t}, \ldots, \sigma_J^{-1}\left(s^0, p^0\right) - x_{Jt}, \pi_1^{-1}\left(s^0, p^0\right) - z_{1t}, \ldots, \pi_J^{-1}\left(s^0, p^0\right) - z_{Jt}\right) |\mathbb{J}(s^0, p^0)|}{f_{\xi,\eta}\left(\sigma_1^{-1}\left(s^0, p^0\right) - x_1^0, \ldots, \sigma_J^{-1}\left(s^0, p^0\right) - x_J^0, \pi_1^{-1}\left(s^0, p^0\right) - z_1^0, \ldots, \pi_J^{-1}\left(s^0, p^0\right) - z_J^0\right) |\mathbb{J}(s^0, p^0)|}$$

at all values of $\left(\sigma_1^{-1}\left(s^0, p^0\right) - x_{1t}, \ldots, \sigma_J^{-1}\left(s^0, p^0\right) - x_{Jt}, \pi_1^{-1}\left(s^0, p^0\right) - z_{1t}, \ldots, \pi_J^{-1}\left(s^0, p^0\right) - z_{Jt}\right)$ traced out by the distribution of $(x_t, z_t)|(s^0, p^0)$. The full support assumption on (x_t, z_t) ensures that this corresponds to *all* values of $(\xi_{1t}, \ldots, \xi_{Jt}, \eta_{1t}, \ldots, \eta_{Jt})$. Noticing that the Jacobian determinants cancel,[19] these ratios then determine the joint density $f_{\xi,\eta}$ up to a multiplicative constant on its full support. This constant, i.e.,

$$f_{\xi,\eta}\left(\sigma_1^{-1}\left(s^0, p^0\right) - x_1^0, \ldots, \sigma_J^{-1}\left(s^0, p^0\right) - x_J^0, \pi_1^{-1}\left(s^0, p^0\right) - z_1^0, \ldots, \pi_J^{-1}\left(s^0, p^0\right) - z_J^0\right)$$

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

is pinned down by the fact that densities must integrate to one. With $f_{\xi,\eta}$ now known, so is each marginal CDF F_{ξ_j} and F_{η_j}. Recalling the normalizations (15), let τ_j^0 denote the known value

$$F_{\xi_j}(0) = F_{\xi_j}\left(\sigma_j^{-1}\left(s^0, p^0\right) - x_j^0\right).$$

Then for any (s_t, p_t), let $x_j(s_t, p_t)$ denote the value of x_{jt} solving

$$\tau_j^0 = F_{\xi_j}\left(\sigma_j^{-1}(s_t, p_t) - x_{jt}\right)$$

so that $\sigma_j^{-1}(s_t, p_t) = -x_j(s_t, p_t)$. Similarly, let υ_j^0 denote the known value

$$F_{\eta_j}(0) = F_{\eta_j}\left(\sigma_j^{-1}\left(s^0, p^0\right) - z_j^0\right).$$

Then for any (s_t, p_t), let $z_j(s_t, p_t)$ denote the value of z_{jt} solving

$$\upsilon_j^0 = F_{\eta_j}\left(\pi_j^{-1}(s_t, p_t) - z_{jt}\right)$$

so that $\rho_j^{-1}(s_t, p_t) = -z_j(s_t, p_t)$. Thus the functions σ_j^{-1} and π_j^{-1} are identified for all j. This implies that all ξ_j (and therefore the demand functions s_j) are identified. □

This provides a constructive proof of the identification of demand and has the added benefit of uniquely determining the supply shocks $\zeta_{1t}, \ldots, \zeta_{Jt}$ as well. However, unlike Theorem 1, here we required a large support condition even for the identification of demand. As before, we can obtain full identification of the random utility model under an additional restriction that utility is quasilinear in the index δ_{jt}.

Theorem 4. *Suppose $x_{jt} \perp\!\!\!\perp \xi_{jt}$ and that Assumptions1, 2′, 3, and 7–10 hold. Then the joint distribution of $(v_{i1t}, \ldots, v_{iJt})$ are identified conditional on any $(\mathcal{J}_t, \{x_{kt}, p_{kt}, \xi_{kt}\}_{k \in \mathcal{J}_t})$ in their support.*

Proof. Under the specification (5), the outside good has market share

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

which is

$$F_\mu\left(-x_{1t} - \xi_{1t}, \ldots, -x_{Jt} - \xi_{Jt}\right). \tag{16}$$

Theorem 3 showed that each ξ_{jt} was identified. The full support assumption (Assumption 9) and (16) then determine F_μ. Since $u_{ijt} = x_{jt} + \xi_{jt} + \mu_j\left(p_{jt}, \omega_{it}\right)$, this gives the result. $\square$

7 Conclusion

We have examined the nonparametric identifiability of a class of models widely used in the empirical literature on demand for differentiated products in industrial organization and a range of other fields of economics. We view the primary message of this work as positive. For many purposes motivating demand estimation, identification of demand is sufficient. Our results show that identification of demand relies primarily on having good instruments. Furthermore, even in extremely rich models of discrete choice with heterogeneous preferences, heteroskedasticity, unobserved choice characteristics, and endogeneity, moving from identification of demand to identification of the full model of random utility requires the same kind of separability and support conditions used to show full identification in even the simplest semiparametric discrete choice models.

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.

Barbera, S., and P. K. Pattanaik (1986): "Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings," *Econometrica*, 54, 707–715.

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.

——— (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 (2008): "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.

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

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

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.

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

Chintagunta, P. (2001): "Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data," *Marketing Science*, 20, 442–456.

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.

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

Falmagne, J.-C. (1978): "A Representation Theorem for Finite Random Scale Systems," *Journal of Mathematical Psychology*, 18, 52–72.

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

Fox, J., and A. Gandhi (2008): "Identifying Selection and Discrete-Continuous Models Using Mixtures," 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 (2007): "What Drives Media Slant," Discussion paper, University of Chicago.

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

Haile, P. A., A. Hortaçsu, and G. Kosenok (2008): "On the Empirical Content of Quantal Response Equilibrium," *American Economic Review*, 98, forthcoming.

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. 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.

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.

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

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

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

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.

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.

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

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

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

Shaked, A., and J. Sutton (1982): "Relaxing Price Competition Through Product Differentiation," *The Review of Economic Studies*, 49(1), 3–13.

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