Identification of BLP-Style Models

Phil Haile
Yale University

Fall 2018

Paper(s)

Berry, S. and P. Haile, "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, September 2014.

also...

Berry, S., A. Gandhi, and P. Haile, "Connected Substitutes and Invertibility of Demand," *Econometrica*, September 2013.

Berry, S. and P. Haile, "Identification of Nonparametric Simultaneous Equations Models with a Residual Index Structure," *Econometrica*, January 2018.

Berry, S. and P. Haile, "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," March 2010.

Berry, S. and P. Haile, "Identification in Differentiated Products Markets," *Annual Reviews in Economics*, September 2016.

Nonparametric Identification: Why Do We Care?

1. How should we think about parametric estimators used in practice? Are the functional form/distributional assumptions merely approximations in finite samples, or essential maintained hypotheses?

Nonparametric Identification: Why Do We Care?

1. How should we think about parametric estimators used in practice? Are the functional form/distributional assumptions merely approximations in finite samples, or essential maintained hypotheses?
2. Guidance for applied work
   - what types of instruments are needed?
   - what are the tradeoffs between
     - IV needs
     - functional form/distributional restrictions
     - better data (e.g., micro data, panel data)

Nonparametric Identification: Why Do We Care?

1. How should we think about parametric estimators used in practice? Are the functional form/distributional assumptions merely approximations in finite samples, or essential maintained hypotheses?
2. Guidance for applied work
   - what types of instruments are needed?
   - what are the tradeoffs between
     - IV needs
     - functional form/distributional restrictions
     - better data (e.g., micro data, panel data)
3. Step toward new estimators.

Prior work...

- taste heterogeneity (semiparametric), **no endogeneity**, e.g.
  - Ichimura & Thompson (1998), Gautier & Kitamura (2009)
  - Briesch, Chintagunta & Matzkin (2005)
- endogeneity w/control function (e.g., Blundell & Powell (2004))
  - supply and demand→fully simultaneous system (same for many economic models with interacting agents or multiple interrelated choices by a single decision maker)
  - Blundell and Matzkin (2014): fully simultaneous systems admit the required triangular structure only under very special functional form restrictions
- semiparametric w/composite error (e.g., Lewbel 2000)
  - $v_{ij} = x_{ij}\beta + \epsilon_{ij} \quad \epsilon_{ij} \sim F\left(\cdot|x_{ij}\right)$
  - consider price elasticity w/heterogeneity and endogeneity
    - heterogeneity: new price or change in product characteristics $\implies$ new distn of ϵ
    - endogeneity: must hold unobservable tastes and product char fixed

A Starting Point: The BLP Demand Model

- Random utility discrete choice

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

$$\beta_{it}^{(k)} = \beta_0^{(k)} + \beta_1^{(k)} d_{it} + \beta_2^{(k)} \nu_{it}^{(k)}$$
$$(\epsilon_{ijt}, \nu_{it}) \perp\!\!\!\perp (x_t, p_t)$$
$$\epsilon_{ijt} \text{ i.i.d. extreme value, } \nu_{it}^{(k)} \text{ i.i.d. normal}$$

Key features:

1. rich preference heterogeneity
2. explicit product/market-specific unobservable ξ_{jt} ($\Longrightarrow$ endogeneity).

How is the BLP demand model identified?

Traditional “informed intuition”

- exogenous changes in choice sets: exclusion restrictions
- supply side may help
- functional form/distributional assumptions?

Examples

Build some intuition from familiar parametric models

1. multinomial logit
2. nested logit
3. mixed logit (BLP)

Multinomial Logit

Consumer i's (conditional indirect) utility from product j in market t

$$v_{ijt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

Notice the linear index in this model:

$$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

Market shares (choice probabilities) are nonlinear functions of the indices:

$$s_{jt} = \frac{e^{\delta_{jt}}}{1 + \sum_k e^{\delta_{kt}}}.$$

Inversion and Instruments

The map from indices to market shares is easily inverted, using the share of the "outside good" 0:

$$\delta_{jt} = ln(s_{jt}) - ln(s_{0t})$$

Inversion and Instruments

The map from indices to market shares is easily inverted, using the share of the "outside good" 0:

$$\delta_{jt} = ln(s_{jt}) - ln(s_{0t})$$

Remembering the definition of the index, we have

$$ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}.$$

Looks like regression model; need instruments for price.

Rewrite

$$ln(s_{jt}) - ln(s_{0t}) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}.$$

Let $x_{jt} = \left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, fix $x_{jt}^{(2)}$ at arbitrary value $\bar{x}_{jt}^{(2)}$, divide through by β_1 (or $|\beta_1|$), and rewrite as

$$x_{jt}^{(1)} + \tilde{\xi}_{jt} = \frac{1}{\beta_1}\left(ln(s_{jt}) - ln(s_{0t})\right) + \frac{\alpha}{\beta}p_{jt}$$

where $\tilde{\xi}_{jt} = \frac{\xi_{jt}}{\beta^{(1)}} + \frac{\bar{x}_{jt}^{(2)}\beta^{(2)}}{\beta^{(1)}}$. LHS equal to a tightly parameterized function of shares and price.

Nested Logit

Still restrictive but adds flexibility through one extra parameter. Inversion of market shares has form

$$\ln(s_{jt}) - \ln(s_{0t}) = x_{jt}\beta - \alpha p_{jt} + (1-\lambda)\ln(s_{j/g,t}) + \xi_{jt}$$

Again looks like a regression equation, but now need instrument for $\ln(s_{j/g,t})$—i.e., for a particular function of the share vector $(s_1, \ldots, s_J)$—not just for p_{jt}.

Nested Logit

Still restrictive but adds flexibility through one extra parameter. Inversion of market shares has form

$$\ln(s_{jt}) - \ln(s_{0t}) = x_{jt}\beta - \alpha p_{jt} + (1-\lambda)\ln(s_{j/g,t}) + \xi_{jt}$$

Again looks like a regression equation, but now need instrument for $\ln(s_{j/g,t})$—i.e., for a particular function of the share vector $(s_1, \ldots, s_J)$—not just for p_{jt}.

Can rewrite as:

$$x_{jt}^{(1)} + \tilde{\xi}_{jt} = \frac{1}{\beta^{(1)}}\left(ln(s_{jt}) - ln(s_{0t}) - (1-\lambda)\ln(s_{j/g,t})\right) + \frac{\alpha}{\beta^{(1)}}p_{jt}$$

Same LHS as before, but RHS now a more complicated function of shares and price.

The "BLP Inversion"

Demand system has inverse, usually written in terms of δ or ξ. If first component of x has a non-random coefficient we can write inverse as

$$x_{jt}^{(1)} + \tilde{\xi}_{jt} = \frac{1}{\beta^{(1)}} \tilde{\delta}_j \left(s_t, p_t, x_t^{(2)}, \theta \right)$$

where the function $\tilde{\delta}_j$ has to be evaluated numerically and depends nonlinearly on params of the random coefficients.

Note: same LHS as before, but RHS now a more complicated function of prices and market shares, *all of which are correlated with* $\tilde{\xi}_t$, i.e., *endogenous.* So we need more IV.

Identification Strategy for Market Level Data

We generalize the approach used in the parametric examples, for a nonparametric generalization of the BLP model:

1. **Index**. Index restriction within a very general nonparametric random utility model.
2. **Inversion**. Generalized multivariate inversion of choice probabilities: express each index as function of endogenous variables.
3. **Instruments**. IV to identify the inverse market share functions and, therefore, the model's structural errors ξ. When those are known, identification of demand is trivial.

Identification, Loosely

The word "identification" is widely used in economics, but often incorrectly/imprecisely. A loose but correct definition is that an *object* is identified if it is uniquely determined by the *population distribution* of observables.

Identification, Loosely

The word "identification" is widely used in economics, but often incorrectly/imprecisely. A loose but correct definition is that an *object* is identified if it is uniquely determined by the *population distribution* of observables.

what objects? e.g.,

- parameter of a hypothesized model, e.g., $y = x\beta + \epsilon$
- a particular average over a distribution of parameters (e.g., individual treatment effects) from a hypothesized model
- other functional (e.g., welfare change from a tax or merger) of a hypothesized model

Note that "model" appears in each example. This is not an accident. One can't even ask the question "is the true quantity XX identified?" without an abstract concept defining "the true quantity XX" This means a model of some kind.

Identification, Formally

Following Hurwicz (1950), Koopmans and Reiersol (1950), Berry and Haile (2018):

- a *structure* S is a data generating process, i.e., a set of probabilistic or functional relationships between the observable and latent variables that implies ("generates") a joint distribution of the observables

Identification, Formally

Following Hurwicz (1950), Koopmans and Reiersol (1950), Berry and Haile (2018):

- a *structure* S is a data generating process, i.e., a set of probabilistic or functional relationships between the observable and latent variables that implies ("generates") a joint distribution of the observables
- let $\mathfrak{S}$ denote the set of all structures; $S_0 \in \mathfrak{S}$ the true structure

Identification, Formally

Following Hurwicz (1950), Koopmans and Reiersol (1950), Berry and Haile (2018):

- a *structure* S is a data generating process, i.e., a set of probabilistic or functional relationships between the observable and latent variables that implies ("generates") a joint distribution of the observables
- let $\mathfrak{S}$ denote the set of all structures; $S_0 \in \mathfrak{S}$ the true structure
- a *hypothesis* is any nonempty subset of $\mathfrak{S}$
- hypothesis $\mathcal{H}$ is true (satisfied) if $S_0 \in \mathcal{H}$

Identification, Formally

Following Hurwicz (1950), Koopmans and Reiersol (1950), Berry and Haile (2018):

- a *structure* S is a data generating process, i.e., a set of probabilistic or functional relationships between the observable and latent variables that implies ("generates") a joint distribution of the observables
- let $\mathfrak{S}$ denote the set of all structures; $S_0 \in \mathfrak{S}$ the true structure
- a *hypothesis* is any nonempty subset of $\mathfrak{S}$
- hypothesis $\mathcal{H}$ is true (satisfied) if $S_0 \in \mathcal{H}$
- a *structural feature* $\theta(S_0)$ is a functional of the true structure

Identification, Formally

Following Hurwicz (1950), Koopmans and Reiersol (1950), Berry and Haile (2018):

- a *structure* S is a data generating process, i.e., a set of probabilistic or functional relationships between the observable and latent variables that implies ("generates") a joint distribution of the observables
- let $\mathfrak{S}$ denote the set of all structures; $S_0 \in \mathfrak{S}$ the true structure
- a *hypothesis* is any nonempty subset of $\mathfrak{S}$
- hypothesis $\mathcal{H}$ is true (satisfied) if $S_0 \in \mathcal{H}$
- a *structural feature* $\theta(S_0)$ is a functional of the true structure

Definition. A structural feature $\theta(S_0)$ is *identified* (or *point identified*, or *identifiable*) under the hypothesis $\mathcal{H}$ if $\theta(S_0)$ is uniquely determined within the set $\{\theta(S) : S \in \mathcal{H}\}$ by the joint distribution of observables. (Equivalently, given $\mathcal{H}$, there are no *observationally equivalent* structures S_1 and S_2 in $\mathcal{H}$ with $\theta(S_1) \neq \theta(S_2)$.)

Remarks on Identification

- *Identification is not even defined without the notion of a true structure within a class defined by maintained hypotheses (what we usually call a "model").* The model may be simple or complicated, may involve economics or only hypothesized statistical relationships (e.g., Rubin causal model). But identification presumes that there are structural features (abstract notions not part of the data themselves) that one wishes to uncover.
- *Identification has nothing to do with a given sample or an estimator.* In fact, strictly speaking it is not even about what one could learn from an infinitely large sample
- *Identification is necessary but not sufficient for existence of suitable estimators*
- *Identification of a structural feature $\theta(S_0)$ may hold even when there are observationally equivalent structures (i.e., when the true structure is itself not identified).*

Our Model, for Market Level Data

Demand Model

- consumer i in market t chooses one product $j \in \mathcal{J}_t$
- condition on any observed consumer characteristics (so treat these fully flexibly)
- "market" $=$ choice set $= \{\mathcal{J}_t, \chi_t\}$
  - $\chi_t =$ matrix of all product & mkt characteristics
- henceforth condition on $\mathcal{J}_t = \mathcal{J} = \{0, 1, \ldots, J\}$.

Market/Product Characteristics

- $x_{jt} \in \mathbb{R}^K$, exogenous observables
- $p_{jt} \in \mathbb{R}$, endogenous observables (price)
- $\xi_{jt} \in \mathbb{R}$, market/choice-specific unobservable (to us)
- let $x_t = (x_{1t}, \ldots, x_{Jt})$, $p_t = (p_{1t}, \ldots, p_{Jt})$, $\xi_t = (\xi_{1t}, \ldots, \xi_{Jt})$

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

Random Utility Model

- conditional indirect utilities ("utilities")

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

- normalized relative to good 0 $\implies v_{i0t} = 0$
- $(v_{i1t}, \ldots, v_{iJt}) \sim F_v(\cdot|\chi_t)$

Random Utility Model

- conditional indirect utilities ("utilities")

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

- normalized relative to good 0 $\implies v_{i0t} = 0$
- $(v_{i1t}, \ldots, v_{iJt}) \sim F_v(\cdot|\chi_t)$
- note:
  - we are <u>not</u> building up $F_v(\cdot|\chi_t)$ from specification of utilty fctns, e.g., $v_{ijt} = v_j(x_t, \xi_t, p_t; \theta)$
  - so far, extremely general random utility model;
    one significant restriction: scalar jt-level unobservable ξ_{jt}.

Index Restriction

- partition $x_{jt} = \left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, $x_{jt}^{(1)} \in \mathbb{R}$
- let $\delta_{jt} = x_{jt}^{(1)}\beta_j + \xi_{jt}, \quad \delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$

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

Index Restriction

- partition $x_{jt} = \left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, $x_{jt}^{(1)} \in \mathbb{R}$
- let $\delta_{jt} = x_{jt}^{(1)}\beta_j + \xi_{jt}, \quad \delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$

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

- recall: satisfied in all the examples
- can relax linearity of index: only monotonicity in ξ_{jt} essential

Index Restriction

- partition $x_{jt} = \left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, $x_{jt}^{(1)} \in \mathbb{R}$
- let $\delta_{jt} = x_{jt}^{(1)}\beta_j + \xi_{jt}, \quad \delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$

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

- recall: satisfied in all the examples
- can relax linearity of index: only monotonicity in ξ_{jt} essential
- scale normalization (i.e., wlog in the model) on each ξ_{jt}:
  $\beta_j = 1\ \forall j$; so now $\delta_{jt} = x_{jt}^{(1)} + \xi_{jt}$

Index Restriction

- partition $x_{jt} = \left(x_{jt}^{(1)}, x_{jt}^{(2)}\right)$, $x_{jt}^{(1)} \in \mathbb{R}$
- let $\delta_{jt} = x_{jt}^{(1)}\beta_j + \xi_{jt}, \quad \delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$

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

- recall: satisfied in all the examples
- can relax linearity of index: only monotonicity in ξ_{jt} essential
- scale normalization (i.e., wlog in the model) on each ξ_{jt}: $\beta_j = 1\ \forall j$; so now $\delta_{jt} = x_{jt}^{(1)} + \xi_{jt}$

Henceforth we condition on $x^{(2)}$ and drop superscripts. The model is then completely general wrt $x^{(2)}$.

Discrete Choice Demand

As usual

- each consumer maximizes utility
- demand (i.e., conditional market shares/choice probabilities):

$$s_{jt} = \sigma_j(\chi_t) = \sigma_j(\delta_t, p_t) = \Pr\left(\arg\max_{j \in \mathcal{J}} \; v_{ijt} = j | \delta_t, p_t\right)$$

- assume $\sigma_j(\delta_t, p_t) > 0 \; \forall \delta_t, p_t$ in their support
  (no relevant goods with zero market share in equilibrium)
- let $s_t = (s_{1t}, \ldots, s_{Jt})$.

Discrete Choice Demand

(continued)

The demand system is

$$s_{jt} = \sigma_j(\chi_t) \qquad j = 1, \ldots, J$$

We observe (s_t, x_t, p_t). If we also observed each ξ_{jt} (or if there were no ξ_{jt}) we would directly observe the functions $(\sigma_1, \ldots, \sigma_J)$: identification of market level demand would be trivial!

Discrete Choice Demand

(continued)

The demand system is

$$s_{jt} = \sigma_j\left(\chi_t\right) \qquad j = 1, \ldots, J$$

We observe (s_t, x_t, p_t). If we also observed each ξ_{jt} (or if there were no ξ_{jt}) we would directly observe the functions $(\sigma_1, \ldots, \sigma_J)$: identification of market level demand would be trivial!

Endogeneity through the unobservables ξ_t is the main challenge to nonparametric identification. More difficult than "usual" here because each s_{jt} (and, in our usual models, each p_{jt}) is a function of <u>all J</u> unobservables $(\xi_{1t}, \ldots, \xi_{Jt})$. As we've mentioned before, our usual ideas about handling endogeneity do not apply when there are multiple structural errors in each equation.

Inversion

Inversion

Generalize the Berry '94 invertibility result
(not a routine to calculate, just existence of inverse)

- inverting demand system = "trick" to obtain representation with 1 structural error per equation (true in Berry94 and here)
- Berry, Gandhi and Haile (2013) show invertibility under a pair of mild conditions we call "connected substitutes."

Connected Substitutes

part (i): weak substitutes

Assumption 2 (part i). For $k \neq j$, $\sigma_k(\delta_t, p_t)$ is weakly decreasing in δ_{jt} [and in $(-p_{jt})$] for all $(\delta_t, p_t) \in \mathbb{R}^{2J}$.

- really just a monotonicity assumption: δ_{jt} is "good" [and price is "bad"] (at least on average)
- e.g., holds if
  - δ_{jt} [and p_{jt}] affect only v_{ijt}, not utilities for other goods
  - v_{ijt} increasing in $(\delta_{jt}, -p_{jt})$

    (both standard, but not necessary).

Connected Substitutes

part (ii): connected strict substitution

Definition 1. *Good j substitutes to good k at* (δ_t, p_t) if $\sigma_k(\delta_t, p_t)$ is strictly decreasing in δ_{jt} and in $(-p_{jt})$.

Connected Substitutes

part (ii): connected strict substitution

Definition 1. *Good j substitutes to good k at (δ_t, p_t)* if $\sigma_k(\delta_t, p_t)$ is strictly decreasing in δ_{jt} and in $(-p_{jt})$.

Represent (strict) substitution with matrix $\Sigma(\delta_t, p_t)$, whoe entries are

$$\Sigma_{j+1,k+1} = \begin{cases} 1\{\text{good } j \text{ substitutes to good } k \text{ at } x\} & j > 0 \\ 0 & j = 0. \end{cases}$$

Connected Substitutes

part (ii): connected strict substitution

Definition 1. *Good j substitutes to good k at* (δ_t, p_t) if $\sigma_k(\delta_t, p_t)$ is strictly decreasing in δ_{jt} and in $(-p_{jt})$.

Represent (strict) substitution with matrix $\Sigma(\delta_t, p_t)$, whoe entries are

$$\Sigma_{j+1,k+1} = \begin{cases} 1\{\text{good } j \text{ substitutes to good } k \text{ at } x\} & j > 0 \\ 0 & j = 0. \end{cases}$$

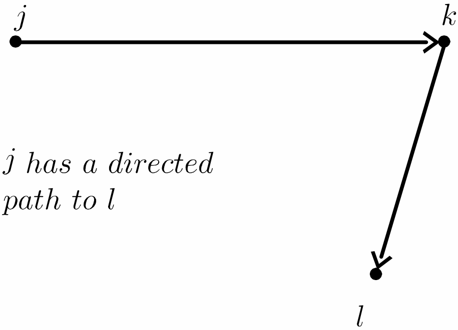
Assumption 2 (part ii). For all (δ_t, p_t) in their support, the directed graph of $\Sigma(\delta_t, p_t)$ has, from every node $k \neq 0$, a directed path to node 0.

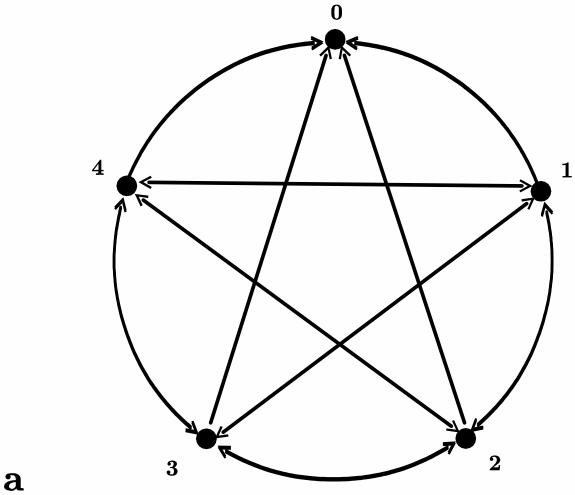

j
k

j *substitutes to* k

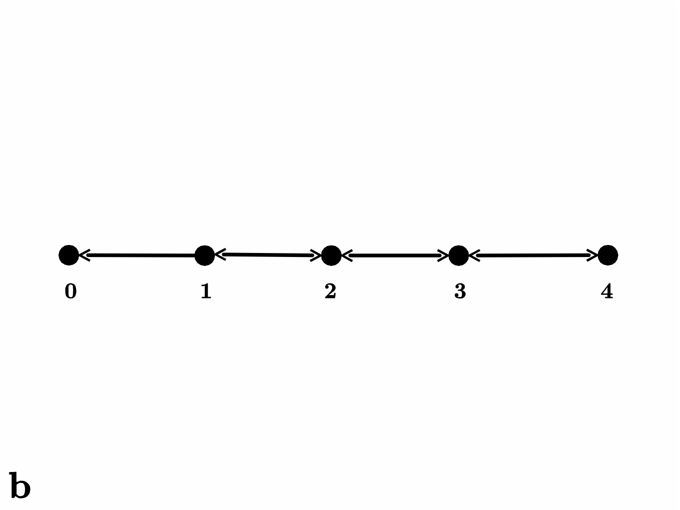
j substitutes to k

k substitutes to j

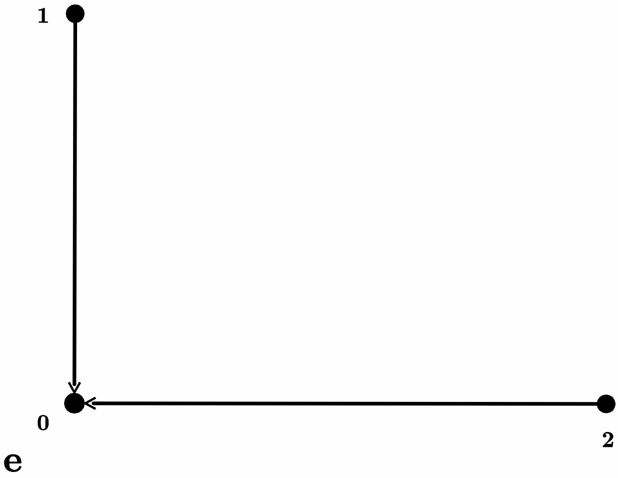

j
k
l
j has a directed path to l



multinomial/nested/mixed logit, probit, etc.



vertical differentiation (e.g., Mussa-Rosen, Bresnahan)



independent goods, with an outside good

Connected Substitutes and Inversion

- CS condition hard to “break” in a discrete choice setting
- equivalent to there being no strict subset of the goods that substitute only among themselves; in other words, all goods actually “belong” in one demand system

Connected Substitutes and Inversion

- CS condition hard to "break" in a discrete choice setting
- equivalent to there being no strict subset of the goods that substitute only among themselves; in other words, all goods actually "belong" in one demand system

Lemma (BGH). Under Assumptions 1-2, for each j there exists a function σ_j^{-1} such that $\delta_{jt} = \sigma_j^{-1}(s_t, p_t)$ for all (s_t, p_t) in their support.

Notes: Existence of each inverse σ_j^{-1}, not characterization. Inversion result actually applies without the discrete choice framework, so room to use this in other kinds of demand systems.

Identification with Market Level Data

Observables

Market level data:

$$M_t, x_t, p_t, s_t, z_t$$

where z_t are excluded IV for demand. Later, add to this list observed cost shifters w_t and instruments y_t excluded from firms' cost functions. Overlap among $(M_t, \mathrm{w}_t, x_t, z_t, y_t)$ is possible.

Using the Inverse

Review of Nonparametric Regression (Newey & Powell, 2003)

Separable Nonparametric Regression Model

- $y_i = \Gamma(x_i) + \epsilon_i$
- IV conditions
  1. "exclusion" : $E[\epsilon_i | z_i] = 0$ a.s.
  2. "completeness": $E[B(x_i) | z_i] = 0$ a.s. implies $B(x_i) = 0$ a.s.

Using the Inverse

Review of Nonparametric Regression (Newey & Powell, 2003)

Separable Nonparametric Regression Model

- $y_i = \Gamma(x_i) + \epsilon_i$
- IV conditions
  1. "exclusion" : $E[\epsilon_i | z_i] = 0$ a.s.
  2. "completeness": $E[B(x_i)|z_i] = 0$ a.s. implies $B(x_i) = 0$ a.s.
     - nonparametric analog of standard rank condition: instruments move endogenous variables around "enough"
     - given 1, 2 is *necessary and sufficient* for identification of Γ
     - so *if you believe there is a nonparametric foundation for IV regression, this is what you believe*.

Link to Inverse Demand

Inverted demand equations similar to regression equations:

- by inversion, $\delta_{jt} = \sigma_j^{-1}(s_t, p_t) \quad \forall jt$, where $\delta_{jt} \equiv x_{jt} + \xi_{jt}$
  $\implies x_{jt} = \sigma_j^{-1}(s_t, p_t) - \xi_{jt} \quad \forall jt$
- vs. nonparametric regression setup
  - not $y_i = \Gamma(\vec{x}_i) + \epsilon_i$
  - but $x_i = \Phi(\vec{y}_i) + \epsilon_i$, with x_i exogenous (and needed as IV)

Link to Inverse Demand

Inverted demand equations similar to regression equations:

- by inversion, $\delta_{jt} = \sigma_j^{-1}(s_t, p_t) \quad \forall jt$, where $\delta_{jt} \equiv x_{jt} + \xi_{jt}$ $\implies x_{jt} = \sigma_j^{-1}(s_t, p_t) - \xi_{jt} \quad \forall jt$
- vs. nonparametric regression setup
  - not $y_i = \Gamma(\vec{x}_i) + \epsilon_i$
  - but $x_i = \Phi(\vec{y}_i) + \epsilon_i$, with x_i exogenous (and needed as IV)
- nonetheless Newey-Powell identification argument extends easily, allowing identification of each $\sigma_j^{-1}(\cdot, \cdot)$

Link to Inverse Demand

Inverted demand equations similar to regression equations:

- by inversion, $\delta_{jt} = \sigma_j^{-1}(s_t, p_t) \quad \forall jt$, where $\delta_{jt} \equiv x_{jt} + \xi_{jt}$
  $\implies x_{jt} = \sigma_j^{-1}(s_t, p_t) - \xi_{jt} \quad \forall jt$
- vs. nonparametric regression setup
  - not $y_i = \Gamma(\vec{x}_i) + \epsilon_i$
  - but $x_i = \Phi(\vec{y}_i) + \epsilon_i$, with x_i exogenous (and needed as IV)
- nonetheless Newey-Powell identification argument extends easily, allowing identification of each $\sigma_j^{-1}(\cdot, \cdot)$
- plug in (s_t, p_t) to reveal each ξ_{jt}
- then identification of σ (demand) follows trivially.

Identification of Demand

Let $z_t = (z_{1t}, \ldots, z_{Jt})$ denote excluded instruments.

Assumption 3. (mean independence) $E[\xi_{jt}|x_t, z_t] = 0$ a.s.

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

Theorem 1. *Under Assumptions 1–4, each σ_j is identified on* supp (δ_t, p_t), and each ξ_{jt} is identified with probability 1.

$$\left.\begin{array}{l}\text{Assumption 1. (index restriction)}\\ \text{Assumption 2. (connected substitutes)}\end{array}\right\} \rightarrow J \text{ inverse share eqns}$$

$$\left.\begin{array}{l}\text{Assumption 3. (mean indep.)}\\ \text{Assumption 4. (completeness)}\end{array}\right\} \text{nonparametric IV regression}$$

Note:

- weaker completeness notions (e.g., Andrews, 2011) generally will do
- extensions in Berry-Haile (2014) for nonseparable index
- constructive arguments (no completeness) in BH (2014, 2017).

*Sketch of proof (mostly Newey-Powell)

- $$\begin{aligned} E\left[\xi_{jt} | x_t, z_t\right] &= E\left[\sigma_j^{-1}\left(s_t, p_t\right) \middle| x_t, z_t\right] - x_{jt} \\ &= 0 \text{ a.s. by mean independence} \end{aligned}$$

*Sketch of proof (mostly Newey-Powell)

- $$E\left[\xi_{jt}|x_t, z_t\right] = E\left[\sigma_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt}$$
  $$= 0 \text{ a.s. by mean independence}$$
- suppose $E\left[\tilde{\sigma}_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt} = 0$ a.s.

*Sketch of proof (mostly Newey-Powell)

- $E\left[\xi_{jt}|x_t, z_t\right] = E\left[\sigma_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt}$
  $= 0 \text{ a.s. by mean independence}$
- suppose $E\left[\tilde{\sigma}_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt} = 0$ a.s.
- $\implies E\left[B\left(s_t, p_t\right)|x_t, z_t\right] = 0 \quad \text{a.s.}$
  - where $B\left(s_t, p_t\right) = \sigma_j^{-1}\left(s_t, p_t\right) - \tilde{\sigma}_j^{-1}\left(s_t, p_t\right)$

*Sketch of proof (mostly Newey-Powell)

- $$E\left[\xi_{jt}|x_t, z_t\right] = E\left[\sigma_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt}$$
  $$= 0 \text{ a.s. by mean independence}$$
- suppose $E\left[\tilde{\sigma}_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt} = 0$ a.s.
- $\implies E\left[B\left(s_t, p_t\right)|x_t, z_t\right] = 0$ a.s.
  - where $B\left(s_t, p_t\right) = \sigma_j^{-1}\left(s_t, p_t\right) - \tilde{\sigma}_j^{-1}\left(s_t, p_t\right)$
- completeness $\implies \tilde{\sigma}_j^{-1} = \sigma_j^{-1}$ a.s.
- $\implies \sigma_j^{-1}$ identified (only the true functions satisfies the IV moments)

*Sketch of proof (mostly Newey-Powell)

- $$E\left[\xi_{jt}|x_t, z_t\right] = E\left[\sigma_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt}$$
  $$= 0 \text{ a.s. by mean independence}$$
- suppose $E\left[\tilde{\sigma}_j^{-1}\left(s_t, p_t\right)\middle| x_t, z_t\right] - x_{jt} = 0$ a.s.
- $\implies E\left[B\left(s_t, p_t\right)|x_t, z_t\right] = 0 \quad$ a.s.
  - where $B\left(s_t, p_t\right) = \sigma_j^{-1}\left(s_t, p_t\right) - \tilde{\sigma}_j^{-1}\left(s_t, p_t\right)$
- completeness $\implies \tilde{\sigma}_j^{-1} = \sigma_j^{-1}$ a.s.
- $\implies \sigma_j^{-1}$ identified (only the true functions satisfies the IV moments)
- $\xi_{jt} = -x_{jt} + \sigma_j^{-1}\left(s_t, p_t\right)$, so each ξ_{jt} a.s.

*Sketch of proof (mostly Newey-Powell)

- $E\left[\xi_{jt}|x_t, z_t\right] = E\left[\sigma_j^{-1}(s_t, p_t)\middle| x_t, z_t\right] - x_{jt}$
  $= 0$ a.s. by mean independence
- suppose $E\left[\tilde{\sigma}_j^{-1}(s_t, p_t)\middle| x_t, z_t\right] - x_{jt} = 0$ a.s.
- $\implies E\left[B(s_t, p_t)|x_t, z_t\right] = 0 \quad$ a.s.
    - where $B(s_t, p_t) = \sigma_j^{-1}(s_t, p_t) - \tilde{\sigma}_j^{-1}(s_t, p_t)$
- completeness $\implies \tilde{\sigma}_j^{-1} = \sigma_j^{-1}$ a.s.
- $\implies \sigma_j^{-1}$ identified (only the true functions satisfies the IV moments)
- $\xi_{jt} = -x_{jt} + \sigma_j^{-1}(s_t, p_t)$, so each ξ_{jt} a.s.
- identification of demand follows: market shares observed and all arguments of $\sigma_j(\delta_t, p_t)$ are now known.

An Interpretation

Obviously no way to avoid need for IV for the endogenous variables.

The model has limited dependent variables, rich heterogeneity, fully simultaneous system (each outcome dependent on the whole vector of structural errors in nonlinear way).

The *best possible hope* is identification of demand under same IV conditions needed for identification of homogeneous regression with a separable scalar error. This holds.

Demand System as the Primitive of Interest

Demand (the functions σ_j) is the main feature of interest on the consumer side of model. It determines all demand counterfactuals (up to extrapolation/interpolation, as usual). Demand is the input from the consumer model needed to estimate marginal costs, test models of supply, etc.

Demand is not necessarily enough to identify changes in *consumer welfare*. As usual, valid notions of aggregate welfare changes can be obtained under an additional assumption of quasilinearity. Welfare changes (changes in CS) then identified whenever demand is.

Discussion: Instruments

Reminder: Our Instruments

1. $x_t \in \mathbb{R}^J$ =a vector of exog product characteristics ("BLP IV")
2. $z_t \in \mathbb{R}^J$ = cost shifters, proxies ("Hausman IV"), Waldfogel IV.

IV Requirements: Exclusion

Why do we need both demand shifters and cost shifters/proxies/etc. even to identify just demand?

IV Requirements: Exclusion

Why do we need both demand shifters and cost shifters/proxies/etc. even to identify just demand?

- each $\sigma_j(\delta_t, p_t)$ depends on a J-vector of structural errors ξ_t; our usual intuition about identification with one IV per endogenous variables does not hold in such cases

IV Requirements: Exclusion

Why do we need both demand shifters and cost shifters/proxies/etc. even to identify just demand?

- each $\sigma_j(\delta_t, p_t)$ depends on a J-vector of structural errors ξ_t; our usual intuition about identification with one IV per endogenous variables does not hold in such cases
- inversion is "trick" to obtain one structural error per equation, so that IV (or simultaneous eqn) arguments can be applied

IV Requirements: Exclusion

Why do we need both demand shifters and cost shifters/proxies/etc. even to identify just demand?

- each $\sigma_j(\delta_t, p_t)$ depends on a J-vector of structural errors ξ_t; our usual intuition about identification with one IV per endogenous variables does not hold in such cases
- inversion is "trick" to obtain one structural error per equation, so that IV (or simultaneous eqn) arguments can be applied
- but each inverse equation has $2J$ endogenous variables:

$$\begin{aligned} \xi_{jt} &= \sigma_j^{-1}(s_t, p_t) - x_{jt} \\ &\quad \text{or, inverting for } s_{jt}, \\ s_{jt} &= h_j(s_{-jt}, p_t, x_{jt}, \xi_{jt}) \end{aligned}$$

  - x_{jt} is IV "for itself"
  - we need $2J - 1$ excluded IV "for" s_{-jt}, p_t.

More Intuition for This

- We want to learn the function $\sigma(\delta, p)$ (including all partial derivatives at all (δ, p) in their support).
- We know we changes in each price, holding all others fixed, and in a way that isn't confounded by changes in ξ. This is a need for instruments that move prices.
- But we also need to hold δ fixed!
- How is this possible if δ is not observable?
- Using the bijection between δ and shares (given p), we can control δ by controlling shares. So we rely on instruments for shares.

The BLP Instruments

Why do we need these in particular?

- consider model with exogenous prices: now condition on p_t

The BLP Instruments

Why do we need these in particular?

- consider model with exogenous prices: now condition on p_t
- demand still depends on J-dimensional ξ_t
- inverse takes the form

$$\begin{aligned} \xi_{jt} &= \sigma_j^{-1}(s_t) - x_{jt} \\ &\text{or} \\ s_{jt} &= h_j\left(s_{-jt}, x_{jt}, \xi_{jt}\right) \end{aligned}$$

The BLP Instruments

Why do we need these in particular?

- consider model with exogenous prices: now condition on p_t
- demand still depends on J-dimensional ξ_t
- inverse takes the form

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

or

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

- we still need $J-1$ instruments for s_{-jt}
  - cost shifters/proxies/Waldfogel IV won't do: *they act only through prices*
  - demand shifters x_{-jt} act on shares through eqm prices and directly through choice problem *They are the only possibility!*

The BLP Instruments

More simply: we have seen that in the most general model we need independent variation in all J shares and all J prices. So no matter how much variation we can generate in prices, we need *something else* that moves all shares at any given price vector. The BLP IV are the only candidates.

The Index Restriction

OK, but why are the BLP instruments excludable?

- recall that we conditioned on $x_t^{(2)}$, so inverse is really $\sigma_j^{-1}\left(s_t, p_t; x_t^{(2)}\right)$
- conditioning on $x_t^{(2)} = x_t^{(2)}$ "instrumenting for itself"
- the index restriction is what leaves $x_t^{(1)}$ out of σ_j^{-1}, making them available as an instrument for shares

So the BLP instruments and the index restriction *both* appear essential, at least without other restrictions or better data. . .

Other Restrictions: e.g., price in the index

Example 1

- suppose $\delta_{jt} = \xi_{jt} - \alpha p_{jt}$, $\alpha = 1$ wlog
  (so price, not $x^{(1)}$, in the index)
- inverse then has form (treating all x fully generally now)

$$\begin{aligned} \xi_{jt} &= \sigma_j^{-1}(s_t) + p_{jt} \\ &\text{or} \\ p_{jt} &= -\sigma_j^{-1}(s_t) + \xi_{jt} \end{aligned}$$

- now need only J instruments (shifting s_t), e.g., cost shifters, proxies, Waldfogel.

Price and X in the Index

Example 2

- suppose $\delta_{jt} = \xi_{jt} + x_{jt}^{(1)}\beta - \alpha p_{jt}$, $\alpha = 1$ wlog
- inverse then has form

$$\begin{aligned} \xi_{jt} &= \sigma_j^{-1}(s_t) - x_{jt}\beta + p_{jt} \\ &\text{or} \\ p_{jt} &= -\sigma_j^{-1}(s_t) + x_{jt}\beta + \xi_{jt} \end{aligned}$$

- now the BLP instruments x_{-jt} can also be used

Many other possibilities, e.g., adding symmetry, exchangeability, nesting, etc.; i.e., many possible tradeoffs between functional form restrictions and IV requirements.

Better Data: Micro Data

We can move the frontier if we observe

- individual choices AND
- consumer-choice-specific observables: $d_{it} = (d_{i1t}, \ldots, d_{iJt})$
  - csr characteristics d_{it}
  - interactions $d_{it} \times x_{jt}$ or $h(d_{it}, x_{jt})$
  - inherently csr-product specific d_{ijt}

$$\left.\begin{array}{l}\text{- interactions } d_{it} \times x_{jt} \text{ or } h(d_{it}, x_{jt}) \\ \text{- inherently csr-product specific } d_{ijt}\end{array}\right\} \text{“}d_{ijt}\text{” — essential}$$

Micro Data in Practice

Examples of d_{ijt} from the literature:

- family size$\times$car size (Goldberg 95, BLP 04, Petrin 02)
- travel time/distance
  - for mode of transport (McFadden, 77)
  - to hospital, school, or retailer (e.g., Capps, Dranove, & Satterthwaite 03, Nielson 18, Burda, Harding & Hausman, 09)
- exposure to product-specific advertising (Ackerberg 01)
- consumer-newspaper political match (Gentzkow-Shapiro, 09)
- household-neighborhood demographic match (Bayer et al. 07).

Random Utility Discrete Choice Model

Similar to before: $(v_{i1t},\ldots,v_{iJt}) \sim F_v(\cdot|\chi_{it})$, where

- $\chi_{it} = (d_{it}, x_t, , p_t, \xi_t)$
- $d_{it} = (d_{i1t},\ldots,d_{iJt})$, $d_{ijt} \in \mathbb{R}$
  (other *ijt*-level observables conditioned out, treated completed generally)
- choice probabilities are

$$s_{ijt} = \sigma_j(d_{it}, x_t, p_t, \xi_t).$$

What Does Micro Data Add?

With micro data, d_{it} provides variation in consumers' choice problems <u>within a market</u>, i.e., without contamination from variation in the unobservables ξ_t. Use this variation instead of IV to learn about "substitution patterns."

Exploiting the Micro Data Variation

Within a market, each d_{it} maps to a new choice probability. Let $\mathcal{D}_t$ be the support of d_{it} (in market t). For each $d \in \mathcal{D}_t$ we observe a choice probability $s_t(d)$. Let

$$\mathcal{S}_t = s_t(\mathcal{D}_t).$$

One "Common Choice Probability"

Assume

- index restriction: $F_v\left(\cdot|\chi_{it}\right) = F_v\left(\cdot|\delta_{it}, x_t, p_t\right)$, where

$$\delta_{ijt} = d_{ijt}\gamma + \xi_{jt}$$

$(|\gamma| = 1 \text{ wlog})$

One "Common Choice Probability"

Assume

- index restriction: $F_v\left(\cdot|\chi_{it}\right) = F_v\left(\cdot|\delta_{it}, x_t, p_t\right)$, where

$$\delta_{ijt} = d_{ijt}\gamma + \xi_{jt}$$

  $(|\gamma| = 1 \text{ wlog})$
- connected substitutes

One "Common Choice Probability"

Assume

- index restriction: $F_v\left(\cdot|\chi_{it}\right) = F_v\left(\cdot|\delta_{it}, x_t, p_t\right)$, where

$$\delta_{ijt} = d_{ijt}\gamma + \xi_{jt}$$

  ($|\gamma| = 1$ wlog)
- connected substitutes
- common choice probability: there exists $\bar{s}$ such that $\bar{s} \in \mathcal{S}_t$ for all t.

Identification with One Common Choice Probability

Theorem (see microdata paper): identification of demand, requiring instruments only for prices.

Intuition:

- invertibility$\longrightarrow$unique d_t^* such that $\sigma\left(d_t^*, x_t, p_t, \xi_t\right) = \bar{s}$
- we can observe this d_t^*
- then we have "regression" equation

$$d_{jt}^* + \xi_{jt} = \sigma^{-1}\left(\bar{s}, x_t, p_t\right).$$

  similar to before (Newey-Powell structure), but now no need to IV for $\bar{s}$ because it is a constant!
- so we need IV only for prices
- identification of σ^{-1} implies identification of each ξ_t.

Identification with One Common Choice Probability

Compared to results with market level data:

- BLP IV not required—need only instruments for prices
- completely general treatment of *all* x_t in the random utility model
- but, requires index restriction involving d_{ijt}
  (but substantially weaker in the paper than what shown here).

Supply

What are some possible objects of interest?

- marginal costs at equilibrium quantities
- marginal cost functions
- marginal cost shocks
- reduced-form for equilibrium prices
- form of oligopoly competition

Adding a Supply Side

Start with known form of oligopoly competition.

Use known demand + known FOC to identify marginal costs, mc fctns.

First-Order Conditions

Assumption 7a. For each $j = 1, \ldots, J$, there exists a known function ψ_j such that in equilibrium

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

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

- this holds under standard supply models, e.g., competitive pricing, monopoly pricing, oligopoly quantity setting or price setting, (all with multi- or single-product firms)
- e.g., in BLP (multi-product price setting) $mc_{jt} = p_{jt} + \Delta_{jt}^{-1} s_t$
- ψ_j "product-specific generalized MR function".

Identification of Marginal Costs

...

...is now immediate given identification of demand:

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

and everything on RHS is known

(this is just the "small miracle" mentioned previously).

Identification of Marginal Cost Functions

Unless one assumes constant mc, we need mc *functions* for counterfactuals that would change equilibrium quantities.

- mc_{jt} is "known," so now we want to identify "regression" model

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

  where

  - Q_{jt} is vector of quantities of all goods produced by firm producing good j
  - w_{jt} and $\omega_{jt} \in \mathbb{R}$ are observed and unobserved cost shifters
  - c_j is an unknown function

- this is standard nonparametric IV regression; we will need IV for Q_{jt}.

Instruments for Firm-Specific Quantities

- typically many of these (overidentification likely...)
  - excluded own demand shifters
  - demand shifters for other goods
  - cost shifters of other firms
  - market-level shifters like population (when properly excluded from our demand model given X, D)
  - (all affect Q_{jt} through equilibrium)
- identification of mc fctns by standard results
  - Chernozhukov-Hansen (2005) with nonseparable ω_{jt}
  - Newey-Powell (2003) with separable ω_{jt}.

Identifying Cost Shocks with Unknown Supply Model

Why?

- allows us to identify *reduced form* for equilibrium prices $p_t = \pi\left(x_t, \xi_t, \mathrm{w}_t, \omega_t\right)$
- useful for discrimination between supply models

Identifying Cost Shocks with Unknown Supply Model

Why?

- allows us to identify *reduced form* for equilibrium prices $p_t = \pi\left(x_t, \xi_t, \mathrm{w}_t, \omega_t\right)$
- useful for discrimination between supply models

How? (details in paper)

- assume prices characterized by (unknown) FOCs for each good
- assume marginal costs have index structure similar to that in demand model
- show invertibility of system of FOCs
- estimate inverted supply system using instruments.

Discrimination Between Models of Firm Conduct

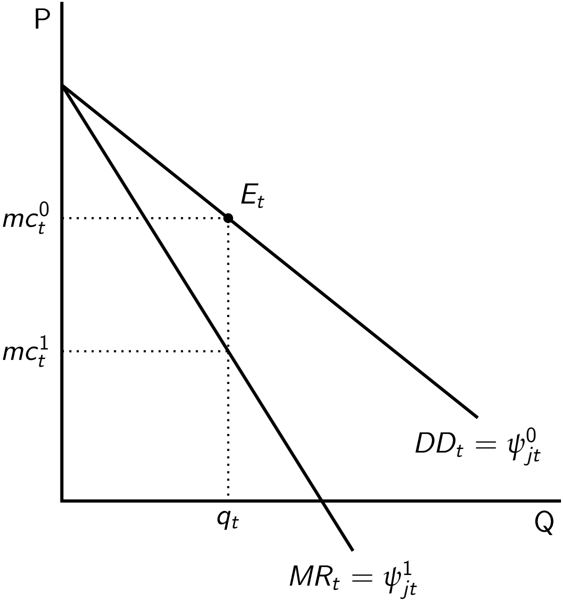
Background

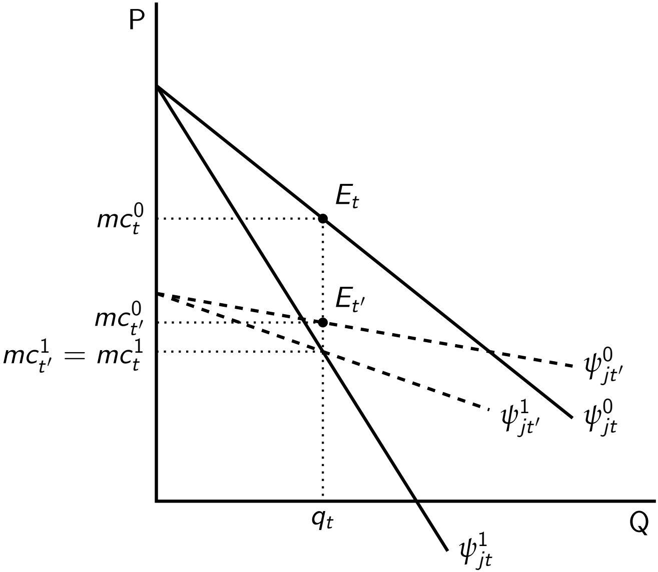
- old literature in IO: (e.g., Bresnahan 1982, 1989)
  - estimate firm "conduct" from market outcomes
  - better interpretation: test candidate models
- *intuition:* estimate demand; then observable "rotations of demand" will change markups—differently in different models–while leaving the equilibrium quantity (and therefore marginal cost) unchanged
- *formal results:* Lau (1982), very limited
  - homogeneous goods, homogeneous firms
  - nonstochastic demand and cost
  - problematic "conjectural variations" framework with "conduct parameter"

Approach 1: Known Cost Shocks

Sketch

- in some results above, we identified the cost shocks without specifying the supply model: work with this case
- known shocks $\iff$ no shocks
- we can then generalize old ideas: rotation of the "product-specific marginal revenue function" under the true model can rule out a false model.


P
E_t
mc_t^0
mc_t^1
$DD_t = \psi_{jt}^0$
q_t
Q
$MR_t = \psi_{jt}^1$


P
Q
E_t
$E_{t'}$
mc_t^0
$mc_{t'}^0$
$mc_{t'}^1 = mc_t^1$
q_t
$\psi_{jt'}^0$
$\psi_{jt'}^1$
ψ_{jt}^0
ψ_{jt}^1

Generalize from the Example

Recall FOC: $mc_j\left(Q_{jt}, \kappa_{jt}\right) = \psi_j\left(s_t, M_t, D_t\left(s_t, p_t\right), p_t\right)$

Generalize from the Example

Recall FOC: $mc_j\left(Q_{jt},\kappa_{jt}\right)=\psi_j\left(s_t,M_t,D_t\left(s_t,p_t\right),p_t\right)$

Theorem 9. Suppose demand identified + cost shocks identified. Oligopoly model is testable: If
(i) $t'\neq t$ and/or $j'\neq j$,
(ii) $mc_j\left(\cdot\right)=mc_{j'}\left(\cdot\right)$, and
(iii) $\left(Q_{jt},\kappa_{jt}\right)=\left(Q_{j't'},\kappa_{j't'}\right)$
then $\psi_j\left(s_t,D_t\left(s_t,p_t\right),p_t\right)=\psi_{j'}\left(s_{t'},D_{t'}\left(s_{t'},p_{t'}\right),p_{t'}\right)$ must hold.

$\implies$ testable restriction of any candidate oligopoly model.

Generalization of Bresnahan/Lau

Theorem shows that Bresnahan's intuition can be formalized for a much more general class of models, although

- key role of "product-specific MR function" ψ_j rather than rotation of demand
- many things can affect ψ_j, not just changes in demand shifters:
  - shocks (ξ_k, ω_k)
  - changes in other firm's characteristics
  - change in number of competitors
  - change in industry ownership structure
  - ...

Open questions: statistical procedure for finite sample.

Approach 2: Unknown Cost Shocks

- with unknown cost shock, 2 observations no longer enough to contradict the FOC $mc_j\left(Q_{jt},\kappa_{jt}\right)=\psi_j\left(s_t,M_t,D_t\left(s_t,p_t\right),p_t\right)$
- but similar restriction applies, averaging over cost shocks (conditional on instruments)
- $\rightarrow$**Theorem 10** in paper.

Conclusions

Identification in BLP-type models turns out to be "boring"

Conclusions

Identification in BLP-type models turns out to be "boring"

- the main requirement is a standard one: adequate instruments
- functional forms mainly play usual roles
  - approximation in finite samples, interpolation, extrapolation
  - filling in for gap between ideal exogenous variation and that available in practice

Conclusions

Identification in BLP-type models turns out to be "boring"

- the main requirement is a standard one: adequate instruments
- functional forms mainly play usual roles
  - approximation in finite samples, interpolation, extrapolation
  - filling in for gap between ideal exogenous variation and that available in practice
- adequate instruments:
  - some tradeoffs between IV needs and functional form restrictions
  - micro data can move this frontier.

Conclusions

Alternative estimation strategies?

- We made progress on identification by focusing directly on the conditional distribution of random utilities and on the demand system. This differs from usual approach of building these up from a parametric specification of individual utilities using random parameters. Here, allowing a *more general* random utility model (or, just focusing directly on demand) actually makes the analysis of identification *simpler*.
- Same strategy may be productive in thinking about new ways to impose flexible structure for estimation. E.g., focus on specification/estimation of inverse demand functions, not utility functions. Parametric/Semi-Parametric/Nonparametric? Gandhi and Nevo (2012), Souza-Rodrigues (2012), Compiani (2018).