1 Why demand analysis/estimation?

There is a huge literature in recent empirical industrial organization which focuses on estimation of demand models. Why??

Demand estimation seems mundane. Indeed, most IO theory concerned about supply-side (firm-side). However, important determinants of firm behavior are **costs**, which are usually unobserved.

For instance, consider a fundamental question in empirical IO: how much market power do firms have? Market power measured by markup: $\frac{p-mc}{p}$. Problem: mc not observed! For example, you observe high prices in an industry. Is this due to market power, or due to high costs? Cannot answer this question directly, because we don't observe costs.

The "new empirical industrial organization" (NEIO; a moniker coined by Bresnahan (1989)) is motivated by this data problem. NEIO takes an *indirect approach*, whereby we obtain estimate of firms' markups by estimating firms' demand functions.

Intuition is most easily seen in monopoly example:

- $\max_p pq(p) - C(q(p))$, where $q(p)$ is demand curve.
- FOC: $q(p) + pq'(p) = C'(q(p))q'(p)$
- At optimal price p^*, **Inverse Elasticity Property** holds:

$$(p^* - MC(q(p^*))) = -\frac{q(p^*)}{q'(p^*)}$$

or

$$\frac{p^* - mc\left(q(p^*)\right)}{p^*} = -\frac{1}{\epsilon(p^*)},$$

where $\epsilon(p^*)$ is $q'(p^*)\dfrac{p^*}{q(p^*)}$, the price elasticity of demand.

- Hence, if we can estimate $\epsilon(p^*)$, we can infer what the markup $\frac{p^*-mc(q(p^*))}{p^*}$ is, even when we don't observe the marginal cost $mc\left(q(p^*)\right)$.
- Similar exercise holds for oligopoly case (as we will show below).
- Caveat: validity of exercise depends crucially on using the right supply-side model (in this case: monopoly without entry possibility).

If costs were observed: markup could be estimated directly, and we could test for vaalidity of monopoly pricing model (ie. test whether markup= $\frac{-1}{\epsilon}$).

In these notes, we begin by reviewing some standard approaches to demand estimation, and motivate why recent literature in empirical IO has developed new methodologies.

2 Review: demand estimation

- Linear demand-supply model:

$$\begin{aligned} \text{Demand:} \quad & q_t^d = \gamma_1 p_t + \mathbf{x}_{\mathbf{t1}}'\beta_{\mathbf{1}} + u_{t1} \\ \text{Supply:} \quad & p_t = \gamma_2 q_t^s + \mathbf{x}_{\mathbf{t2}}'\beta_{\mathbf{2}} + u_{t2} \\ \text{Equilibrium:} \quad & q_t^d = q_t^s \end{aligned}$$

- Demand function summarizes consumer preferences; supply function summarizes firms' cost structure
- Focus on estimating demand function:

$$\text{Demand: } q_t = \gamma_1 p_t + \mathbf{x}_{\mathbf{t1}}'\beta_{\mathbf{1}} + u_{t1}$$

- If u_1 correlated with u_2, then p_t is endogenous in demand function: cannot estimate using OLS. Important problem.

- Instrumental variable (IV) methods: assume there are instruments Z's so that $E(u_1 \cdot \mathbf{Z}) = 0$.
- Properties of appropriate instrument Z for endogenous variable p:
  1. Uncorrelated with error term in demand equation: $E(u_1 Z) = 0$. **Exclusion** restriction. (order condition)
  2. Correlated with endogenous variable: $E(Zp) \neq 0$. (rank condition)
- The x's are exogenous variables which can serve as instruments:
  1. x_{t2} are *cost shifters*; affect production costs. Correlated with p_t but not with u_{t1}: use as instruments in demand function.
  2. x_{t1} are *demand shifters*; affect willingness-to-pay, but not a firm's production costs. Correlated with q_t but not with u_{2t}: use as instruments in supply function.

The demand models used in empirical IO different in flavor from "traditional" demand specifications. Start by briefly showing traditional approach, then motivating why that approach doesn't work for many of the markets that we are interested in.

2.1 "Traditional" approach to demand estimation

- Consider modeling demand for two goods 1,2 (Example: food and clothing).
- Data on prices and quantities of these two goods across consumers, across markets, or over time.
- Consumer demand determined by utility maximization problem:

$$\max_{x_1,x_2} U(x_1, x_2) \text{ s.t. } p_1 x_2 + p_2 x_2 = M$$

- This yields demand functions $x_1^*(p_1, p_2, M)$, $x_2^*(p_1, p_2, M)$.

- Equivalently, start out with *indirect utility function*

$$V(p_1, p_2, M) = U(x_1^*(p_1, p_2, M), x_2^*(p_1, p_2, M))$$

- Demand functions derived via *Roy's Identity*:

$$x_1^*(p_1, p_2, M) = -\frac{\partial V}{\partial p_1} / \frac{\partial V}{\partial M}$$
$$x_2^*(p_1, p_2, M) = -\frac{\partial V}{\partial p_2} / \frac{\partial V}{\partial M}$$

  This approach is often more convenient empirically.

- This "standard" approach not convenient for many markets which we are interested in: automobile, airlines, cereals, toothpaste, etc. These markets characterized by:

  - Many alternatives: too many parameters to estimate using traditional approach
  - At individual level, usually only choose one of the available options (discrete choices). Consumer demand function not characterized by FOC of utility maximization problem.

  These problems have been addressed by

  - Modeling demand for a product as demand for the characteristics of that product: **Hedonic** analysis (Rosen (1974), Bajari and Benkard (2005)).

    This can be difficult in practice when there are many characteristics, and characteristics not continuous.
  - Discrete choice: assume each consumer can choose at most *one* of the available alternatives on each purchase occasion. This is the approach taken in the moden empirical IO literature.

3 Discrete-choice approach to modeling demand

- Starting point: McFadden's ((1978),(1981)) *random utility* framework.

- There are J alternatives $j = 1, \ldots, J$. Each purchase occasion, each consumer i divides her income y_i on (at most) one of the alternatives, and on an "outside good":

$$\max_{j,z} U_i(x_j, z) \text{ s.t. } p_j + p_z z = y_i$$

where

  - x_j are chars of brand j, and p_j the price
  - z is quantity of outside good, and p_z its price
  - outside good ($j = 0$) denotes the non-purchase of any alternative (that is, spending entire income on other types of goods).

- Substitute in the budget constraint ($z = \frac{y-p_n}{p_z}$) to derive *conditional indirect utility functions* for each brand:

$$U^*_{ij} = U_i(x_j, \frac{y - p_j}{p_z}).$$

If outside good is bought:

$$U^*_{i0} = U_i(0, \frac{y}{p_z}).$$

- Consumer chooses the brand yielding the highest cond. indirect utility:

$$\max_j U^*_{ij}$$

- U^*_{ij} specified as sum of two parts. The first part is a function $V_{ij}(\cdots)$ of the observed variables (prices, characteristics, etc.). The second part is a "utility shock", consisting of choice-affecting elements not observed by the econometrician:

$$U^*_{ij} = V_{ij}(p_j, p_z, y_i) + \epsilon_{ij}$$

The utility shock ϵ_{ij} is observed by agent i, not by econometrician: we call this a **structural error**. From agent's point of view, utility and choice are *deterministic*.

- Given this specification, the probability that consumer i buys brand j is:

$$D_{ij} = \text{Prob}\left\{\epsilon_{i0}, \ldots, \epsilon_{iJ} : U^*_{ij} > U^*_{ij'} \text{ for } j' \neq j\right\}$$

  If households are identical, so that $V_{ij} = V_{i'j}$ for i, i', and $\vec{\epsilon} \equiv \{\epsilon_{i0}, \ldots, \epsilon_{iJ}\}'$ is *iid* across agents i (and there are a very large number of agents), then D_{ij} is also the *aggregate market share.*

- Hence, specific distributional assumptions on $\vec{\epsilon}$ determine the functional form of choice probabilities. Two common distributional assumptions are:

  1. $(\epsilon_{i0}, \ldots, \epsilon_{iJ})$ distributed multivariate normal: **multinomial probit**. Choice probabilities do not have closed form, but they can be simulated (Keane (1994), McFadden (1989)). (cf. GHK simulator, which we describe in a different set of lecture notes.)

     But model becomes awkward when there are large number of choices, because number of parameters in the variance matrix Σ also grows very large.

  2. $(\epsilon_{ij},\ j = 0, \ldots, J)$ distributed *i.i.d.* type I extreme value across i:

$$F(\epsilon) = \exp\left[-\exp\left(-\frac{\epsilon - \eta}{\mu}\right)\right]$$

     with the location parameter $\eta = 0.577$ (Euler's constant), and the scale parameter (usually) $\mu = 1$.

     This leads to multinomial logit choice probabilities:

$$D_{ij}(\cdots) = \frac{\exp(V_{ij})}{\sum_{j'=1,\ldots,J} \exp(V_{ij'})}$$

     Normalize $V_0 = 0$. (Because $\sum_{j=1}^{J} D_{ij} = 1$ by construction.)

     Convenient, tractable form for choice probabilities, which scales easily when the number of goods increases. For this reason, the multinomial logit model is basis for many demand papers in empirical IO.

Problems with multinomial logit

Despite its tractibility, the MNL model has restrictive implications, which are particularly unattractiveness for its uss in a demand setting. Specifically: the odds ratio between any two brands j, j' doesn't depend on number of alternatives available

$$\frac{D_j}{D_{j'}} = \frac{\exp(V_j)}{\exp(V_{j'})}$$

This is the **Independence of Irrelevant Alternatives (IIA)** property.

Example: Red bus/blue bus problem:

- Assume that city has two transportation schemes: walk, and red bus, with shares 50%, 50%. So odds ratio of walk/RB= 1.
- Now consider introduction of third option: train. IIA implies that odds ratio between walk/red bus is still 1. Unrealistic: if train substitutes more with bus than walking, then new shares could be walk 45%, RB 30%, train 25%, then odds ratio walk/RB=1.5.
- What if third option were blue bus? IIA implies that odds ratio between walk/red bus would still be 1. Unrealistic: BB is perfect substitute for RB, so that new shares are walk 50%, RB 25%, bb 25%, and odds ratio walk/RB=2!
- So this is especially troubling if you want to use logit model to predict penetration of new products.

Implication: invariant to introduction (or elimination) of some alternatives.

If interpret D_{ij} as market share, IIA implies restrictive substitution patterns:

$$\varepsilon_{a,c} = \varepsilon_{b,c}, \text{ for all brands } a, b \neq c.$$

If $V_j = \beta_j + \alpha(y - p_j)$, then $\varepsilon_{a,c} = -\alpha p_c D_c$, for all $c \neq a$: Price decrease in brand a attracts proportionate chunk of demand from all other brands. Unrealistic!

Because of this, the multinomial logit model has been "tweaked" in order to eliminate the implications of IIA:

1. Nested logit: assume particular correlation structure among $(\epsilon_{i0}, \ldots, \epsilon_{iN})$. Within-nest brands are "closer substitutes" than across-nest brands (This model is generated by assuming that the utility shocks $\vec{\epsilon}$ follow a "generalized extreme value" distribution, cf. Maddala (1983, chap. 2)). See Goldberg (1995) for an application of this to automobile demand.

   (Diagram of demand structure from Goldberg paper. One shortcoming of this approach is that the researcher must know the "tree structure" of the model.)

2. Random coefficients: assume logit model, but for agent i:

$$U^*_{ij} = X'_j \beta_i - \alpha_i p_j + \epsilon_{ij}$$

   (coefficients are agent-specific). This allows for valuations of characteristics, and price-sensitivities, to vary across households. But note that, unlike nested logit model, IIA is still present at the individual-level decision-making problem here; the individual-level choice probability is still multionomial-logit in form:

$$D_{ij} = \frac{\exp\left(X'_j \beta_i - \alpha_i p_j\right)}{\sum_{j'} \exp\left(X'_{j'} \beta_i - \alpha_i p_{j'}\right)}.$$

   But aggregate market share is

$$\int \frac{\exp\left(X'_j \beta_i - \alpha_i p_j\right)}{\sum_{j'} \exp\left(X'_{j'} \beta_i - \alpha_i p_{j'}\right)} \cdot dF(\alpha_i, \beta_i)$$

   and differs from individual choice probability. At the aggregate level, IIA property disappears.

   We will focus on this model below, because it has been much used in the recent literature.

4 Berry (1994) approach to estimate demand in differentiated product markets

Methodology for estimating differentiated-product discrete-choice demand models, using aggregate data.

Data structure: *cross-section* of market shares:

j	$\hat{s}_j$	p_j	X_1	X_2
A	25%	\$1.50	red	large
B	30%	\$2.00	blue	small
C	45%	\$2.50	green	large

Total market size: M
J brands

Note: this is different data structure than that considered in previous contexts: here, all variation is across brands (and no variation across time or markets).

Background: Trajtenberg (1989) study of demand for CAT scanners. Disturbing finding: coefficient on price is *positive*, implying that people prefer more expensive machines! Upward-sloping demand curves.

(Tables of results from Trajtenberg paper)

Possible explanation: quality differentials across products not adequately controlled for. In equilibrium of a diff'd product market where each product is valued on the basis of its characteristics, brands with highly-desired characteristics (higher quality) command higher prices. Unobserved quality leads to price endogeneity.

■■■

Here, we start out with simplest setup, with most restrictive assumptions, and later describe more complicated extensions.

Derive market-level share expression from model of discrete-choice at the individual household level (i indexes household, j is brand):

$$U_{ij} = \underbrace{X_j\beta - \alpha p_j + \xi_j}_{\equiv \delta_j} + \epsilon_{ij}$$

where we call δ_j the "mean utility" for brand j (the part of brand j's utility which is common across all households i).

■■■

Econometrician observes neither ξ_j or ϵ_{ij}, but household i observes both: these are both "structural errors".

$\xi_1, \ldots, \xi_J$ are interpreted as "unobserved quality". All else equal, consumers more willing to pay for brands for which ξ_j is high.

Important: ξ_j, as unobserved quality, is correlated with price p_j (and also potentially with characteristics X_j). It is the source of the endogeneity problem in this demand model.

Make logit assumption that $\epsilon_{ij} \sim \; iid$ TIEV, across consumers i and brands j.

Define choice indicators:

$$y_{ij} = \begin{cases} 1 & \text{if } i \text{ chooses brand } j \\ 0 & \text{otherwise} \end{cases}$$

Given these assumptions, choice probabilities take MN logit form:

$$Pr\left(y_{ij} = 1 | \beta, x_{j'}, \xi_{j'}, j' = 1, \ldots, J\right) = \frac{\exp\left(\delta_j\right)}{\sum_{j'=0}^{J} \exp\left(\delta_{j'}\right)}.$$

Aggregate market shares are:

$$\begin{aligned} s_j &= \frac{1}{M}\left[M \cdot Pr\left(y_{ij} = 1 | \beta, x_{j'}, \xi_{j'}, j' = 1, \ldots, J\right)\right] = \frac{\exp\left(\delta_j\right)}{\sum_{j'=1}^{J} \exp\left(\delta_{j'}\right)} \\ &\equiv \tilde{s}_j\left(\delta_0, \delta_1, \ldots, \delta_J\right) \equiv \tilde{s}_j\left(\alpha, \beta, \xi_1, \ldots, \xi_J\right). \end{aligned}$$

$\tilde{s}(\cdots)$ is the "predicted share" function, for fixed values of the parameters α and β, and the unobservables $\xi_1, \ldots, \xi_J$.

■■■

- Data contains observed shares: denote by $\hat{s}_j, j = 1, \ldots, J$

  (Share of outside good is just $\hat{s}_0 = 1 - \sum_{j=1}^{J} \hat{s}_j$.)

- Model + parameters give you predicted shares: $\tilde{s}_j(\alpha, \beta, \xi_1, \ldots, \xi_J), j = 1, \ldots, J$

- Principle: Estimate parameters α, β by finding those values which "match" observed shares to predicted shares: find α, β so that $\tilde{s}_j(\alpha, \beta)$ is as close to $\hat{s}_j$ as possible, for $j = 1, \ldots, J$.

- How to do this? Note that you cannot do **nonlinear least squares**, i.e.

$$\min_{\alpha,\beta} \sum_{j=1}^{J} \left(\hat{s}_j - \tilde{s}_j(\alpha, \beta, \xi_1, \ldots, \xi_J)\right)^2 \tag{1}$$

  This problem doesn't fit into standard NLS framework, because you need to know the ξ's to compute the predicted share, and they are not observed.

■■■

Berry (1994) suggests a clever IV-based estimation approach.

Assume there exist instruments Z so that $E(\xi Z) = 0$

Sample analog of this moment condition is

$$\frac{1}{J} \sum_{j=1}^{J} \xi_j Z_j = \frac{1}{J} \sum_{j=1}^{J} (\delta_j - X_j \beta + \alpha p_j) Z_j$$

which converges (as $J \to \infty$) to zero at the true values α_0, β_0. We wish then to estimate (α, β) by minimizing the sample moment conditions.

Problem with estimating: we do not know δ_j! Berry suggest a *two-step approach*

First step: Inversion

- If we equate $\hat{s}_j$ to $\tilde{s}_j(\delta_0, \delta_1, \ldots, \delta_J)$, for all j, and normalize $\delta_0 = 0$, we get a system of J nonlinear equations in the J unknowns $\delta_1, \ldots, \delta_J$:

$$\begin{aligned}\hat{s}_1 &= \tilde{s}_1(\delta_1, \ldots, \delta_J)\\ \vdots \quad & \quad \vdots \\ \hat{s}_J &= \tilde{s}_J(\delta_1, \ldots, \delta_J)\end{aligned}$$

- You can "invert" this system of equations to solve for $\delta_1, \ldots, \delta_J$ as a function of the observed $\hat{s}_1, \ldots, \hat{s}_J$.
- Note: the outside good is $j = 0$. Since $1 = \sum_{j=0}^{J} \hat{s}_j$ by construction, you normalize $\delta_0 = 0$.
- Output from this step: $\hat{\delta}_j \equiv \delta_j(\hat{s}_1, \ldots, \hat{s}_J)$, $j = 1, \ldots, J$ (J numbers)

Second step: IV estimation

- Going back to definition of δ_j's:

$$\begin{aligned}\delta_1 &= X_1\beta - \alpha p_1 + \xi_1\\ \vdots \quad & \quad \vdots \\ \delta_J &= X_J\beta - \alpha p_J + \xi_J\end{aligned}$$

- Now, using estimated $\hat{\delta}_j$'s, you can calculate sample moment condition:

$$\frac{1}{J}\sum_{j=1}^{J}\left(\hat{\delta}_j - X_j\beta + \alpha p_j\right) Z_j$$

  and solve for α, β which minimizes this expression.

- If δ_j is linear in X, p and ξ (as here), then linear IV methods are applicable here. For example, in 2SLS, you regress p_j on Z_j in first stage, to obtain fitted prices $\hat{p}(Z_j)$. Then in second stage, you regression δ_j on X_j and $\hat{p}(Z_j)$.

  Later, we will consider the substantially more complicated case when the underlying demand model is the random-coefficients logit model, as in Berry, Levinsohn, and Pakes (1995).

■■■

What are appropriate instruments (Berry, p. 249)?

- Usual demand case: cost shifters. But since we have cross-sectional (across brands) data, we require instruments to very across brands in a market.
- Take the example of automobiles. In traditional approach, one natural cost shifter could be wages in Michigan.
- But here it doesn't work, because its the same across all car brands (specifically, if you ran 2SLS with wages in Michigan as the IV, first stage regression of price p_j on wage would yield the same predicted price for all brands).
- BLP exploit competition within market to derive instruments. They use IV's like: characteristics of cars of competing manufacturers. Intuition: oligopolistic competition makes firm j set p_j as a function of characteristics of cars produced by firms $i \neq j$ (e.g. GM's price for the Hum-Vee will depend on how closely substitutable a Jeep is with a Hum-Vee). However, characteristics of rival cars should not affect households' valuation of firm j's car.
- In multiproduct context, similar argument for using characteristics of all other cars produced by same manufacturer as IV.
- With panel dataset, where prices and market shares for same products are observed across many markets, could also use prices of product j in other markets as instrument for price of product j in market t (eg. Nevo (2001), Hausman (1996)).

■■■

One simple case of inversion step:

MNL case: predicted share $\tilde{s}_j\left(\delta_1, \ldots, \delta_J\right) = \frac{\exp(\delta_j)}{1+\sum_{j'=1}^{J} \exp\left(\delta_{j'}\right)}$

The system of equations from matching actual to predicted shares is:

$$\hat{s}_0 = \frac{1}{1 + \sum_{j=1}^{J} \exp(\delta_j)}$$
$$\hat{s}_1 = \frac{\exp(\delta_1)}{1 + \sum_{j=1}^{J} \exp(\delta_j)}$$
$$\vdots \qquad \vdots$$
$$\hat{s}_J = \frac{\exp(\delta_J)}{1 + \sum_{j=1}^{J} \exp(\delta_j)}.$$

Taking logs, we get system of linear equations for δ_j's:

$$\log \hat{s}_1 = \delta_1 - \log(\text{denom})$$
$$\vdots \qquad \vdots$$
$$\log \hat{s}_J = \delta_J - \log(\text{denom})$$
$$\log \hat{s}_0 = 0 - \log(\text{denom})$$

which yield

$$\delta_j = \log \hat{s}_j - \log \hat{s}_0, \quad j = 1, \ldots, J.$$

So in second step, run IV regression of

$$(\log \hat{s}_j - \log \hat{s}_0) = X_j\beta - \alpha p_j + \xi_j. \tag{2}$$

Eq. (2) is called a "logistic regression" by bio-statisticians, who use this logistic transformation to model "grouped" data. So in the simplest MNL logit, the estimation method can be described as "logistic IV regression".

See Berry paper for additional examples (nested logit, vertical differentiation).

■■■

4.1 Measuring market power: recovering markups

- Next, we show how demand estimates can be used to derive estimates of firms' markups (as in monopoly example from the beginning).
- From our demand estimation, we have estimated the demand function for brand j, which we denote as follows:

$$D^j\left(\underbrace{X_1,\ldots,X_J}_{\equiv\vec{X}};\underbrace{p_1,\ldots,p_J}_{\equiv\vec{p}};\underbrace{\xi_1,\ldots,\xi_J}_{\equiv\vec{\xi}}\right)$$

- Specify costs of producing brand j:

$$C^j\left(q_j, w_j, \omega_j\right)$$

  where q_j is total production of brand j, w_j are observed cost components associated with brand j (e.g. could be characteristics of brand j), ω_j are unobserved cost components (another structural error)

- Then profits for brand j are:

$$\Pi_j = D^j\left(\vec{X},\vec{p},\vec{\xi}\right)p_j - C^j\left(D^j\left(\vec{X},\vec{p},\vec{\xi}\right), w_j, \omega_j\right)$$

- For multiproduct firm: assume that firm k produces all brands $j \in \mathcal{K}$. Then its profits are

$$\tilde{\Pi}_k = \sum_{j\in\mathcal{K}}\Pi_j = \sum_{j\in\mathcal{K}}\left[D^j\left(\vec{X},\vec{p},\vec{\xi}\right)p_j - C^j\left(D^j\left(\vec{X},\vec{p},\vec{\xi}\right), w_j, \omega_j\right)\right].$$

  Importantly, we assume that there are no (dis-)economies of scope, so that production costs are simply additive across car models, for a multiproduct firm.

- In order to proceed, we need to assume a particular model of oligopolistic competition.

  The most common assumption is *Bertrand (price) competition.* (Note that because firms produce differentiated products, Bertrand solution does not result in marginal cost pricing.)

- Under price competition, equilibrium prices are characterized by J equations (which are the J pricing first-order conditions for the J brands):

$$\frac{\partial \tilde{\Pi}_k}{\partial p_j} = 0, \ \forall j \in \mathcal{K}, \ \forall k$$
$$\Leftrightarrow D^j + \sum_{j' \in \mathcal{K}} \frac{\partial D^{j'}}{\partial p_j} \left(p_{j'} - C_1^{j'}|_{q_{j'} = D^{j'}} \right) = 0$$

where C_1^j denotes the derivative of C^j with respect to first argument (which is the marginal cost function).

- Note that because we have already estimated the demand side, the demand functions $D^j, \ j = 1, \ldots, J$ and full set of demand slopes $\frac{\partial D^{j'}}{\partial p_j}, \ \forall j, j' = 1, \ldots, J$ can be calculated.

  Hence, from these J equations, we can solve for the J margins $p_j - C_1^j$. In fact, the system of equations is linear, so the solution of the marginal costs C_1^j is just

$$\vec{c} = \vec{p} + (\Delta D)^{-1} \vec{D}$$

  where c and D denote the J-vector of marginal costs and demands, and the derivative matrix ΔD is a $J \times J$ matrix where

$$\Delta D_{(i,j)} = \begin{cases} \frac{\partial D^i}{\partial p_j} & \text{if models } (i,j) \text{ produced by the same firm} \\ 0 & \textit{otherwise.} \end{cases}$$

  The markups measures can then be obtained as $\frac{p_j - C_1^j}{p_j}$.

  This is the oligopolistic equivalent of using the "inverse-elasticity" condition to calculate a monopolist's market power.

5 Berry, Levinsohn, and Pakes (1995): Demand and supply estimation using random-coefficients logit model

Next we discuss the case of the random coefficients logit model, which is the main topic of Berry, Levinsohn, and Pakes (1995).

- Assume that utility function is:

  $$u_{ij} = X_j\beta_i - \alpha_i p_j + \xi_j + \epsilon_{ij}$$

  The difference here is that the slope coefficients (α_i, β_i) are allowed to vary across households i.

- We assume that, across the population of households, the slope coefficients (α_i, β_i) are i.i.d. random variables. The most common assumption is that these random variables are jointly normally distributed:

  $$(\alpha_i, \beta_i)' \sim N\left((\bar{\alpha}, \bar{\beta})', \Sigma\right).$$

  For this reason, α_i and β_i are called "random coefficients".

  Hence, $\bar{\alpha}$, $\bar{\beta}$, and Σ are additional parameters to be estimated.

- Given these assumptions, the mean utility δ_j is $X_j\bar{\beta} - \bar{\alpha}p_j + \xi_j$, and

  $$u_{ij} = \delta_j + \epsilon_{ij} + (\beta_i - \bar{\beta})X_j - (\alpha_i - \bar{\alpha})p_j$$

  so that, even if the ϵ_{ij}'s are still i.i.d. TIEV, the composite error is not. Here, the simple MNL inversion method will not work.

- The estimation methodology for this case is developed in Berry, Levinsohn, and Pakes (1995).

- First note: for a given α_i, β_i, the choice probabilities for household i take MNL form:

  $$Pr(i,j) = \frac{\exp\left(X_j\beta_i - \alpha_i p_j + \xi_j\right)}{1 + \sum_{j'=1}^{J}\exp\left(X_{j'}\beta_i - \alpha_i p_{j'} + \xi_{j'}\right)}.$$

- In the whole population, the aggregate market share is just

$$
\begin{aligned}
\tilde{s}_j &= \int \int Pr(i, j,) dG(\alpha_i, \beta_i) \\
&= \int \int \frac{\exp\left(X_j \beta_i - \alpha_i p_j + \xi_j\right)}{1 + \sum_{j'=1}^{J} \exp\left(X_{j'} \beta_i - \alpha_i p_{j'} + \xi_{j'}\right)} dG(\alpha_i, \beta_i) \\
&= \int \int \frac{\exp\left(\delta_j + (\beta_i - \bar{\beta}) X_j - (\alpha_i - \bar{\alpha}) p_j\right)}{1 + \sum_{j'=1}^{J} \exp\left(\delta_{j'} + (\beta_i - \bar{\beta}) X_{j'} - (\alpha_i - \bar{\alpha}) p_{j'}\right)} dG(\alpha_i, \beta_i) \\
&\equiv \tilde{s}_j^{RC}\left(\delta_1, \ldots, \delta_J; \bar{\alpha}, \bar{\beta}, \Sigma\right)
\end{aligned}
\tag{3}
$$

  that is, roughly speaking, the weighted sum (where the weights are given by the probability distribution of (α, β)) of $Pr(i, j)$ across all households.

  The last equation in the display above makes explicit that the predicted market share is not only a function of the mean utilities $\delta_1, \ldots, \delta_J$ (as before), but also functions of the parameters $\bar{\alpha}, \bar{\beta}, \Sigma$. Hence, the inversion step described before will not work, because the J equations matching observed to predicted shares have more than J unknowns (i.e. $\delta_1, \ldots, \delta_J; \bar{\alpha}, \bar{\beta}, \Sigma$).

  Moreover, the expression in Eq. (3) is difficult to compute, because it is a multidimensional integral. BLP propose *simulation methods* to compute this integral. We will discuss simulation methods later. For the rest of these notes, we assume that we can compute $\tilde{s}_j^{RC}$ for every set of parameters $\bar{\alpha}, \bar{\beta}, \Sigma$.

■■■

We would like to proceed, as before, to estimate via GMM, exploiting the population moment restriction $E\left(\xi Z_m\right) = 0, \; i = 1, \ldots, M$. Let $\theta \equiv (\bar{\alpha}, \bar{\beta}, \Sigma)$. Then the sample moment conditions are:

$$
m_{m,J}(\theta) \equiv \frac{1}{J} \sum_{j=1}^{J} \left(\delta_j - X_j \bar{\beta} + \bar{\alpha} p_j\right) Z_{mj}
$$

and we estimate θ by minimizing a quadratic norm in these sample moment functions:

$$
\min_{\theta} Q_J(\theta) \equiv [m_{m,J}(\theta)]'_m W_J [m_{m,J}(\theta)]_m
$$

W_J is a $(M \times M)$-dimensional weighting matrix.

But problem is that we cannot perform inversion step as before, so that we cannot derive $\delta_1, \ldots, \delta_J$.

So BLP propose a "nested" estimation algorithm, with an "inner loop" nested within an "outer loop"

- In the **outer loop**, we iterate over different values of the parameters. Let $\hat{\theta}$ be the current values of the parameters being considered.
- In the **inner loop**, for the given parameter values $\hat{\theta}$, we wish to evaluate the objective function $Q(\hat{\theta})$. In order to do this we must:
  1. At current $\hat{\theta}$, we solve for the mean utilities $\delta_1(\hat{\theta}), \ldots, \delta_J(\hat{\theta})$ to solve the system of equations

  $$\begin{aligned} \hat{s}_1 &= \tilde{s}_1^{RC}\left(\delta_1, \ldots, \delta_J; \hat{\theta}\right) \\ \vdots \quad & \quad \vdots \\ \hat{s}_J &= \tilde{s}_J^{RC}\left(\delta_1, \ldots, \delta_J; \hat{\theta}\right). \end{aligned}$$

     Note that, since we take the parameters $\hat{\theta}$ as given, this system is J equations in the J unknowns $\delta_1(\hat{\theta}), \ldots, \delta_J(\hat{\theta})$.
  2. For the resulting $\delta_1(\hat{\theta}), \ldots, \delta_J(\hat{\theta})$, calculate

  $$Q(\hat{\theta}) = [m_{m,J}(\hat{\theta})]'_m W_J [m_{m,J}(\hat{\theta})]_m \qquad (4)$$

- Then we return to the outer loop, which searches until it finds parameter values $\hat{\theta}$ which minimize Eq. (4).
- Essentially, the original inversion step is now nested inside of the estimation routine.

Note that, typically, for identification, a necessary condition is that:

$$M = dim(\vec{Z}) \geq dim(\theta) > dim(\alpha, \beta) = dim([X, p]).$$

This is because there are coefficeints Σ associated with the distribution of random coefficients. This implies that, even if there were no price endogeneity problem, so that (X, p) are valid instruments, we still need additional instruments in order to identify the additional parameters.[1]

5.1 Estimating equilibrium: both demand and supply side

Within this nested estimation procedure, we can also add a supply side to the RC model.

Let us make the further assumption that marginal costs are constant, and linear in cost components:

$$C_1^j = c^j \equiv w_j \gamma + \omega_j$$

(where γ are parameters in the marginal cost function) then the best-response equations become

$$D^j + \sum_{j' \in K} \frac{\partial D^{j'}}{\partial p_j} \left(p_{j'} - c^j \right) = 0. \tag{5}$$

Assume you have instruments U_j such that $E(\omega U) = 0$. From the discussion previously, these instruments would be "demand shifters" which would affect pricing and sales but unrelated to production costs.

With both demand and supply-side moment conditions, the objective function becomes:

$$Q(\theta, \gamma) = G_J(\theta, \gamma)' W_J G_J(\theta, \gamma)$$

[1]See Moon, Shum, and Weidner (2012).

where G_J is the $(M+N)$-dimensional vector of stacked sample moment conditions:

$$G_J(\theta,\gamma) \equiv \begin{bmatrix} \frac{1}{J}\sum_{j=1}^{J}\left(\delta_j(\theta) - X_j\bar{\beta} + \bar{\alpha}p_j\right) z_{1j} \\ \vdots \\ \frac{1}{J}\sum_{j=1}^{J}\left(\delta_j(\theta) - X_j\bar{\beta} + \bar{\alpha}p_j\right) z_{Mj} \\ \frac{1}{J}\sum_{j=1}^{J}\left(c_j(\theta) - w_j\gamma\right) u_{1j} \\ \vdots \\ \frac{1}{J}\sum_{j=1}^{J}\left(c_j(\theta) - w_j\gamma\right) u_{Nj} \end{bmatrix}$$

where M is the number of demand side IV's, and N the number of supply-side IV's. (Assuming $M + N \geq dim(\theta) + dim(\gamma)$)

The only change in the estimation routine described in the previous section is that the inner loop is more complicated:

In the **inner loop**, for the given parameter values $\hat{\theta}$ and $\hat{\gamma}$, we wish to evaluate the objective function $Q(\hat{\theta},\hat{\gamma})$. In order to do this we must:

1. At current $\hat{\theta}$, we solve for the mean utilities $\delta_1(\hat{\theta}),\ldots,\delta_J(\hat{\theta})$ as previously.
2. For the resulting $\delta_1(\hat{\theta}),\ldots,\delta_J(\hat{\theta})$, calculate

   $$\vec{\tilde{s}}_j^{RC}(\hat{\theta}) \equiv \left(\tilde{s}_1^{RC}(\delta(\hat{\theta})),\ldots,\tilde{s}_J^{RC}(\delta(\hat{\theta}))\right)'$$

   and also the partial derivative matrix

   $$\mathbf{D}(\hat{\theta}) = \begin{pmatrix} \frac{\partial \tilde{s}_1^{RC}(\delta(\hat{\theta}))}{\partial p_1} & \frac{\partial \tilde{s}_1^{RC}(\delta(\hat{\theta}))}{\partial p_2} & \cdots & \frac{\partial \tilde{s}_1^{RC}(\delta(\hat{\theta}))}{\partial p_J} \\ \frac{\partial \tilde{s}_2^{RC}(\delta(\hat{\theta}))}{\partial p_1} & \frac{\partial \tilde{s}_2^{RC}(\delta(\hat{\theta}))}{\partial p_2} & \cdots & \frac{\partial \tilde{s}_2^{RC}(\delta(\hat{\theta}))}{\partial p_J} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \tilde{s}_J^{RC}(\delta(\hat{\theta}))}{\partial p_1} & \frac{\partial \tilde{s}_J^{RC}(\delta(\hat{\theta}))}{\partial p_2} & \cdots & \frac{\partial \tilde{s}_J^{RC}(\delta(\hat{\theta}))}{\partial p_J} \end{pmatrix}$$

   For MN logit case, these derivatives are:

   $$\frac{\partial s_j}{\partial p_k} = \begin{cases} -\alpha s_j(1-s_j) & \text{for } j = k \\ -\alpha s_j s_k & \text{for } j \neq k. \end{cases}$$

3. Use the supply-side best response equations to solve for $c_1(\hat{\theta}), \ldots, c_J(\hat{\theta})$:

$$\vec{s}_j^{RC}(\hat{\theta}) + \mathbf{D}(\hat{\theta}) * \begin{pmatrix} p_1 - c^1 \\ \vdots \\ p_J - c^J \end{pmatrix} = 0.$$

4. So now, you can compute $G(\hat{\theta}, \hat{\gamma})$.

5.2 Simulating the integral in Eq. (3)

The principle of simulation: approximate an expectation as a sample average. Validity is ensured by law of large numbers.

In the case of Eq. (3), note that the integral there is an expectation:

$$\mathcal{E}\left(\bar{\alpha}, \bar{\beta}, \Sigma\right) \equiv E_G \left[\frac{\exp\left(\delta_j + (\beta_i - \bar{\beta})X_j - (\alpha_i - \bar{\alpha})p_j\right)}{1 + \sum_{j'=1}^{J} \exp\left(\delta_{j'} + (\beta_i - \bar{\beta})X_{j'} - (\alpha_i - \bar{\alpha})p_{j'}\right)} | \bar{\alpha}, \bar{\beta}, \Sigma \right]$$

where the random variables are α_i and β_i, which we assume to be drawn from the multivariate normal distribution $N\left((\bar{\alpha}, \bar{\beta})', \Sigma\right)$.

For $s = 1, \ldots, S$ simulation draws:

1. Draw u_1^s, u_2^s independently from N(0,1).
2. For the current parameter estimates $\hat{\bar{\alpha}}$, $\hat{\bar{\beta}}$, $\hat{\Sigma}$, transform $(u_1^s,\ u_2^s)$ into a draw from $N\left((\hat{\bar{\alpha}}, \hat{\bar{\beta}})', \hat{\Sigma}\right)$ using the transformation

$$\begin{pmatrix} \alpha^s \\ \beta^s \end{pmatrix} = \begin{pmatrix} \hat{\bar{\alpha}} \\ \hat{\bar{\beta}} \end{pmatrix} + \hat{\Sigma}^{1/2} \begin{pmatrix} u_1^s \\ u_2^s \end{pmatrix}$$

   where $\hat{\Sigma}^{1/2}$ is shorthand for the "Cholesky factorization" of the matrix $\hat{\Sigma}$. The Cholesky factorization of a square symmetric matrix $\mathbf{\Gamma}$ is the triangular matrix $\mathbf{G}$ such that $\mathbf{G}'\mathbf{G} = \mathbf{\Gamma}$, so roughly it can be thought of as a matrix-analogue of "square root". We use the *lower triangular* version of $\hat{\Sigma}^{1/2}$.

Then approximate the integral by the sample average (over all the simulation draws)

$$\mathcal{E}\left(\hat{\hat{\alpha}}, \hat{\hat{\beta}}, \hat{\Sigma}\right) \approx \frac{1}{S}\sum_{s=1}^{S} \frac{\exp\left(\delta_j + (\beta^s - \hat{\hat{\beta}})X_j - (\alpha^s - \hat{\hat{\alpha}})p_j\right)}{1 + \sum_{j'=1}^{J} \exp\left(\delta_{j'} + (\beta^s - \hat{\hat{\beta}})X_{j'} - (\alpha^s - \hat{\hat{\alpha}})p_{j'}\right)}.$$

For given $\hat{\hat{\alpha}}, \hat{\hat{\beta}}, \hat{\Sigma}$, the law of large numbers ensure that this approximation is accurate as $S \to \infty$.

(Results: marginal costs and markups from BLP paper)

6 Applications

Applications of this methodology have been voluminous. Here discuss just a few.

1. evaluation of VERs In Berry, Levinsohn, and Pakes (1999), this methodology is applied to evaluate the effects of voluntary export restraints (VERs). These were voluntary quotas that the Japanese auto manufacturers abided by which restricted their exports to the United States during the 1980's.

The VERs do not affect the demand-side, but only the supply-side. Namely, firm profits are given by:

$$\pi_k = \sum_{j \in \mathcal{K}} (p_j - c_j - \lambda VER_k) D^j.$$

In the above, VER_k are dummy variables for whether firm k is subject to VER (so whether firm k is Japanese firm). VER is modelled as an "implicit tax", with $\lambda \geq 0$ functioning as a per-unit tax: if $\lambda = 0$, then the VER has no effect on behavior, while $\lambda > 0$ implies that VER is having an effect similar to increase in marginal cost c_j. The coefficient λ is an additional parameter to be estimated, on the supply-side.

Results (effects of VER on firm profits and consumer welfare)

2. Welfare from new goods, and merger evaluation After cost function parameters γ are estimated, you can simulate equilibrium prices under alternative market structures, such as mergers, or entry (or exit) of firms or goods. These counterfactual prices are valid assuming that consumer preferences and firms' cost functions don't change as market structures change. Petrin (2002) presents consumer welfare benefits from introduction of the minivan, and Nevo (2001) presents merger simulation results for the ready-to-eat cereal industry.

3. Geographic differentiation In our description of BLP model, we assume that all consumer heterogeneity is unobserved. Some models have considered types of consumer heterogeneity where the marginal distribution of the heterogeneity in the population is observed. In BLP's original paper, they include household income in the utility functions, and integrate out over the population income distribution (from the Current Population Survey) in simulating the predicted market shares.

Another important example of this type of obbserved consumer heterogeneity is consumers' location. The idea is that the products are geographically differentiated, so that consumers might prefer choices which are located closer to their home. Assume you want to model competition among movie theaters, as in Davis (2006). The utility of consumer i from theater j is:

$$U_{ij} = -\alpha p_j + \beta(L_i - L_j) + \xi_j + \epsilon_{ij}$$

where $(L_i - L_j)$ denotes the geographic distance between the locations of consumer I and theater j. The predicted market shares for each theater can be calculated by integrating out over the marginal empirical population density (ie. integrating over the distribution of L_i). See also Thomadsen (2005) for a model of the fast-food industry, and Houde (2012) for retail gasoline markets. The latter paper is noteworthy because instead of integrating over the marginal distribution of where people live, Houde integrates over the distribution of commuting routes. He argues that consumers are probably more sensitive to a gasoline station's location relative to their driving routes, rather than relative to their homes.

A Additional details: general presentation of random utility models

Introduce the *social surplus function*

$$H(\vec{U}) \equiv \mathbb{E}\left\{\max_{j\in\mathcal{J}}(U_j + \epsilon_j)\right\}$$

where the expectation is taken over some joint distribution of $(\epsilon_1, \ldots, \epsilon_J)$.

For each $\lambda \in [0,1]$, for all values of $\vec{\epsilon}$, and for any two vectors $\vec{U}$ and $\vec{U'}$, we have

$$\max_j(\lambda U_j + (1-\lambda)U'_j + \epsilon_j) \leq \lambda \max_j(U_j + \epsilon_j) + (1-\lambda)\max_j(U'_j + \epsilon_j).$$

Since this holds for all vectors $\vec{\epsilon}$, it also holds in expectation, so that

$$H(\lambda\vec{U} + (1-\lambda)\vec{U'}) \leq \lambda H(\vec{U}) + (1-\lambda)H(\vec{U'}).$$

That is, $H(\cdot)$ is a convex function. We consider its *Fenchel-Legendre transformation*[2] defined as

$$H^*(\vec{p}) = \max_{\vec{U}}(\vec{p}\cdot\vec{U} - H(\vec{U}))$$

where $\vec{p}$ is some J-dimensional vector of choice probabilities. Because H is convex we have that the FOCs characterizing H^* are

$$\vec{p} = \nabla_{\vec{U}} H(\vec{U}). \tag{6}$$

Note that for discrete-choice models, this function is many-to-one. For any constant k, $H(\vec{U}+k) = H(\vec{U}) + k$, and hence if $\vec{U}$ satisfies $\vec{p} = \nabla_{\vec{U}} H(\vec{U})$, then also $\vec{p} = \nabla_{\vec{U}} H(\vec{U}+k)$.

$H^*(\cdot)$ is also called the "conjugate" function of $H(\cdot)$. Furthermore, it turns out that the conjugate function of $H^*(\vec{p})$ is just $H(\vec{U})$ – for this reason, the functions H^* and H have a dual relationship, and

$$H(\vec{U}) = \max_{\vec{p}}(\vec{p}\cdot\vec{U} - H^*(\vec{p})).$$

[2]See Gelfand and Fomin (1965), Rockafellar (1971), Chiong, Galichon, and Shum (2013).

with

$$\vec{U} \in \partial_{\vec{p}} H^*(\vec{p}) \tag{7}$$

where $\partial_{\vec{p}} H^*(\vec{p})$ denotes the *subdifferential* (or, synonymously, subgradiant or subderivative) of H^* at $\vec{p}$. For discrete choice models, this is typically a multi-valued mapping (a correspondence) because $\nabla H(\vec{U})$ is many-to-one.[3] In the discrete choice literature, equation (6) is called the William-Daly-Zachary theorem, and analogous to the Shepard/Hotelling lemmas, for the random utility model. Eq. (7) is a precise statement of the "inverse mapping" from choice probabilities to utilities for discrete choice models, and thus reformulates (and is a more general statement of) the "inversion" result in Berry (1994) and BLP (1995).

For specific assumptions on the joint distribution of $\vec{\epsilon}$ (as with the generalized extreme value case above), we can derive a closed form for the social surplus function $H(\vec{U})$, which immediately yield the choice probabilities via Eq. (6) above.

For the multinomial logit model, we know that

$$H(\vec{U}) = \log\left(\sum_{i=0}^{K} \exp(U_i)\right).$$

From the conjugacy relation, we know that $\vec{p} = \nabla H(\vec{U})$. Normalizing $U_0 = 0$, this leads to $U_i = \log(p_i/p_0)$ for $i = 1, \ldots, K$. Plugging this back into the definition of $H^*(\vec{p})$, we get that

$$H^*(\vec{p}) = \sum_{i'=0}^{K} p_{i'} \log(p_{i'}/p_0) - \log\left(\frac{1}{p_0}\sum_{i'=0}^{K} p_{i'}\right) \tag{8}$$

$$= \sum_{i'=1}^{K} p_{i'} \log p_{i'} - \log p_0 \sum_{i'=1}^{K} p_{i'} + \log p_0 \tag{9}$$

$$= \sum_{i'=0}^{K} p_{i'} \log p_{i'}. \tag{10}$$

[3]Indeed, in the special case where $\nabla H(\cdot)$ is one-to-one, then we have $\vec{U} = (\nabla H(\vec{p}))$. This is the case of the classical Legendre transform.

To confirm, we again use the conjugacy relation $\vec{U} = \nabla H^*(\vec{p})$ to get (for $i = 0, 1, \ldots, K$) that $U_i = \log p_i$. Then imposing the normalization $U_0 = 0$, we get that $U_i = \log(p_i/p_0)$.

References

BAJARI, P., AND L. BENKARD (2005): "Demand Estimation With Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, 113, 1239–1276.

BERRY, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy," *American Economic Review*, 89, 400–430.

BRESNAHAN, T. (1989): "Empirical Studies of Industries with Market Power," in *Handbook of Industrial Organization*, ed. by R. Schmalensee, and R. Willig, vol. 2. North-Holland.

CHIONG, K., A. GALICHON, AND M. SHUM (2013): "Duality in Dynamic Discrete Choice Models," mimeo, Caltech.

DAVIS, P. (2006): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, pp. 964–982.

GELFAND, I., AND S. FOMIN (1965): *Calculus of Variations*. Dover.

GOLDBERG, P. (1995): "Product Differentiation and Oligopoly in International Markets: The Case of the US Automobile Industry," *Econometrica*, 63, 891–951.

HAUSMAN, J. (1996): "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. by T. Bresnahan, and R. Gordon, pp. 209–237. University of Chicago Press.

HOUDE, J. (2012): "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, 102, 2147–2182.

KEANE, M. (1994): "A Computationally Practical Simulation Estimator for Panel Data," *Econometrica*, 62, 95–116.

MADDALA, G. S. (1983): *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.

McFadden, D. (1978): "Modelling the Choice of Residential Location," in *Spatial Interaction Theory and Residential Location*, ed. by A. K. et. al. North Holland Pub. Co.

——— (1981): "Statistical Models for Discrete Panel Data," in *Econometric Models of Probabilistic Choice*, ed. by C. Manski, and D. McFadden. MIT Press.

——— (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 57, 995–1026.

Moon, R., M. Shum, and M. Weidner (2012): "Estimation of Random Coefficients Logit Demand Models with Interactive Fixed Effects," manuscript, University of Southern California.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-eat Cereals Industry," *Econometrica*, 69, 307–342.

Petrin, A. (2002): "Quantifying the Benefits of New Products: the Case of the Minivan," *Journal of Political Economy*, 110, 705–729.

Rockafellar, T. (1971): *Convex Analysis*. Princeton University Press.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34–55.

Thomadsen, R. (2005): "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *RAND Journal of Economics*, pp. 908–929.

Trajtenberg, M. (1989): "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, 97(2), 444–479.