Estimating Models of Entry and Differentiated Products

Isabelle Perrigne, Quang Vuong and Weichen Yan

Rice University and New York University

September 2018

Motivation

- In oligopolies with differentiated products, firms make several decisions:

  1. which products to offer
  2. how to differentiate products (attributes)
  3. how to price offered products

Berry, Levinsohn and Pakes or BLP (1995) focus on pricing with discrete choice model and IV for price endogeneity

Fan (2013) and Ackerberg, Crawford and Hahn (2011) endogenize product characteristics in a BLP setting

Li, Mazur, Park, Roberts, Sweeting and Zhang (2018) and Ciliberto, Murry and Tamer (2016) combine entry model with BLP in airline industry but computationally heavy

- In parallel, literature on entry models using reduced-form profits.

  See e.g. Bresnahan and Reiss (1991), Berry (1992), Seim (2006), Sweeting (2009)

- Empirical motivation: movie release, product positioning, airline companies, merger analysis, etc

Objectives

- A model combining entry and pricing for differentiated products under incomplete information
- A new and computationally friendly estimator for BLP style models with endogenous choice set

Main Ideas

- Entry Model leads to firms' selection through the product qualities and cost shocks

  → Bias in the estimation of the demand and cost parameters in BLP pricing model

  → Brand effects not identified

- Control function approach for firms' selection and IV for price endogeneity

  → Sieves for the correction terms

Related Econometric Literature: Heckman (1979), Robinson (1988), Ahn and Powell (1993),

Das, Newey and Vella or DNV (2003)

Roadmap

1. The Model
2. Identification
3. Estimation
4. Monte Carlo Study

1. The Model

NOTATIONS

$\overline{C} \equiv \{0, 1, 2, ..., \overline{J}\}$ grand set of possible alternatives

"0": outside option j: index for product

X_j characteristics for product j

ξ_j unobserved product heterogeneity $\sim$ product quality, $E(\xi_j) = \mu_j$ 'brand effect'

ϵ_{ij} random component capturing consumer's i taste for product j

P_j price of product j

C actual choice set (including outside option) $\{0\} \subseteq C \subseteq \overline{C}$

W_j cost shifters

ω_j cost shock

κ_j entry cost for product j

BLP Benchmark Setup

Consumer i's utility for product j: $\quad U_{ij} = X_j'(\beta + \sigma \circ \upsilon_i^\beta) - \alpha P_j + \xi_j + \epsilon_{ij}$

$\rightarrow \sigma \circ \upsilon_i^\beta$ random coefficients on X_j

with component $\beta_{ik} \equiv \beta_k + \sigma_k \nu_{ik}^\beta$, (β_k, σ_k) mean and STD of β_{ik}

Outside option $U_{i0} = \epsilon_{i0}$

Set of available products $C = \{0\} \cup C^*$ with characteristics X_{C^*}, prices P_{C^*} and qualities ξ_{C^*}

Discrete choice model: Consumer i chooses product j iff $U_{ij} \geq \max_{j' \in C} U_{ij'}$

Market share for product j: $\quad S_{j:C} = s_{j:C}(X_{C^*}, P_{C^*}, \xi_{C^*}; \theta)$ with $\theta \equiv (\beta', \alpha', \sigma')'$

Firm's marginal cost for product j: $\quad \log MC_j = W_j'\gamma + \omega_j$

Product qualities ξ_j and cost shocks ω_j potentially dependent

Bertrand model of competition to determine the prices P_j

Remark: Econometric model given later

Endogenous Choice Set and Entry

Two-stage game: 1. Firms' Entry determining C^*

2. Pricing (Bertrand competition)

Assumptions: (i) $C^* = E$ set of entering firms (single-product firms)

(ii) Private information in entry game: ξ_j known to firm j and $(\xi_{-j}, \omega_j, \omega_{-j})$ unknown to firm j

(iii) All information revealed in pricing game

(iv) Distributions $F_{\xi\omega|Z}(\cdot,\cdot|Z)$ and $F_{\nu^\beta\epsilon}(\cdot,\cdot)$ common knowledge

$Z = (X, W, R)$ denoting all exogenous variables including market ones

Timing: 1) Firms draw their qualities $\xi = (\xi_1, \ldots, \xi_{\overline{J}})$ from $F_{\xi\omega|Z}(\cdot,\cdot|Z)$ and decide to enter based on their own ξ_j thereby determining the J_E entering firms and products set C

2) The J_E firms draw their cost shocks $\omega_E = \{\omega_j; j \in E\}$ given their qualities $\xi_E = \{\xi_j; j \in E\}$ from $F_{\omega_E|\xi_E Z}(\cdot|\xi_E, Z)$. All information is revealed. Firms decide their prices

Remark: $\omega_E \perp \xi_{-E}|(\xi_E, Z)$ following 2) (could be released)

Perfect Bayesian Equilibrium

Pricing Decisions: $P_E \equiv \{P_{j:E}; j \in E\}$ oligopoly prices solutions of

$$(P_{j:E} - MC_j)\frac{\partial s_{j:C}(X_E, P_E, \xi_E; \theta)}{\partial P_j} + s_{j:C}(X_E, P_E, \xi_E; \theta) = 0$$

where $MC_j = mc_j(W_j, \omega_j; \gamma)$

$\rightarrow P_{j:E} = P_{j:E}(X_E, W_E, \xi_E, \omega_E; \theta, \gamma)$

$\rightarrow$ Firm's j profit $\Pi^*_{j:E} = \pi^*_{j:E}(X_E, W_E, \xi_E, \omega_E; \theta, \gamma)$

Entry Decisions: Entry strategy $D_j = 1$ iff $\mathrm{E}[\Pi_j|\xi_j, Z, D_j = 1] \geq 0$ leading to

$$\sum_{E:j\in E} \mathrm{E}\left[\Pi^*_{j:E} \mid \xi_j, \mathbf{Z}, \mathcal{E} = E\right] \times \Pr[\mathcal{E} = E \mid \xi_j, \mathbf{Z}, D_j = 1] \geq \kappa_j$$

where the expectation is with respect to $(\xi_{E\setminus\{j\}}, \omega_E)$

Monotome pure Bayesian Nash strategies: $D_j = 1\!\!I[\xi_j \geq \xi^*_j(Z)]$ where equilibrium thresholds $\xi^*_j(Z; \theta, \gamma, \kappa)$, $j = 1, \ldots, \overline{J}$ solve the system of equations

$$\begin{aligned}\kappa_j \;&=\; \sum_{E:j\in E} \mathrm{E}\left[\pi^*_{j:E}(X_E, W_E, \xi^*_j, \xi_{E\setminus\{j\}}, \omega_E; \theta, \gamma)|\xi_j = \xi^*_j, \xi_{E\setminus\{j\}} \geq \xi^*_{E\setminus\{j\}}, \xi_{-E} < \xi^*_{-E}, Z\right] \\ &\quad \times \Pr[\xi_{E\setminus\{j\}} \geq \xi^*_{E\setminus\{j\}}, \xi_{-E} < \xi^*_{-E}|\xi_j = \xi^*_j, Z]\end{aligned}$$

$\rightarrow$ Thresholds $\xi^* = \xi^*(Z; \theta, \gamma, \kappa)$

2. Identification

Econometric Setup

Observations: (C, S, P, Z) for each market (omitting index m)

Parameters: $(\alpha, \beta, \gamma, \kappa, \sigma)$ and brand effects $\mu \equiv \{\mu_j; j \in \overline{C}^*\}$

Econometric Model

Demand equation from Berry (1994) inversion

$$\psi_{j:C}(S;\sigma) = X_j\beta - \alpha P_j + \xi_j$$

for $j \in C^*$, where $S \equiv \{S_j; j \in C^*\}$

Supply equation from Bertrand competition

$$\log\left[P_j + \frac{S_j}{\partial s_{j:C}(X_E, P_E, \xi_E; \theta)/\partial P_j}\right] = W_j'\gamma + \omega_j$$

for $j \in E$ ($C^* = E$ because of single-product firms to simplify)

Well-known endogeneity issue in the demand with price P_j correlated with quality ξ_j

→ Use of instruments such as X_j, W_j as well as (X_{-j}, W_{-j})

But also selection because of endogenous choice set!

Recall that firms enter iff drawing higher ξ, i.e. $\xi_j \geq \xi_j^*(Z)$

→ $\mathrm{E}(\xi_j | X_j, C)$ no longer a constant and $\mathrm{E}(\xi_j | X_j, C) \geq \mu_j$

→ $\mathrm{E}(\omega_j | W_j, C)$ no longer a constant either because of the dependence between ω_j and ξ_j

→ Inconsistency of usual estimators

→ We need to account for endogenous choice set or equivalently firms' selection

Firms' Selection

Assumption: $(\xi, \omega) \perp Z$

Comment: Can be somewhat relaxed to conditional independence but allows additive correction terms in demand and supply equations

How to correct for firms' selection?

$\rightarrow$ Control Function Approach

Heckman (1979)

DNV (2003): Extension of the Heckman correction in nonparametric regression with endogenous variables and multivariate selection

But P_j is determined structurally and some assumption of DNV will not be satisfied

Entry probability $\pi_j = \pi_j(Z) \equiv \Pr(D_j = 1|Z)$

Basic Idea: The expectation of functions of $(X_E, W_E, \xi_E, \omega_E)$ conditional on Z and E can be expressed as functions of (X_E, W_E, π) (Extension of DNV (2003))

Controlling for selection in demand, cost and price

$$\begin{aligned} \mathrm{E}(\xi_j|Z,E) &= \mu_j + \lambda^{\xi}_{j:E}[\pi(Z)] \\ \mathrm{E}(\omega_j|Z,E) &= \lambda^{\omega}_{j:E}[\pi(Z)] \\ \mathrm{E}(P_j|Z,E) &= \lambda^{P}_{j:E}[X_E, W_E, \pi(Z)] \equiv \tilde{P}_{j:E} \end{aligned}$$

Comments:

(i) $\mathrm{E}(\xi_j|Z,E)$ and $\mathrm{E}(\omega_j|Z,E)$ depend on Z through $\pi(Z)$ only

(ii) $\pi(Z)$ involve the entry probabilities of all firms, i.e. entering and nonentering

(iii) The functions $\lambda^{\xi}_{j:E}(\cdot)$, $\lambda^{\omega}_{j:E}$ and $\lambda^{P}_{j:E}$ vary with j and E

Identification

Demand Parameters $\theta = (\beta', \alpha, \sigma')'$

Fix pair (j, E) with $j \in E$, conditioning demand equation on (Z, E) gives the partial linear regression model

$$\psi_{j:C}(S;\sigma) = X_j'\beta - \alpha \tilde{P}_{j:E} + \mu_j + \lambda^{\xi}_{j:E}(\pi) + \eta_{j:E}$$

with $\mathrm{E}(\eta_{j:E}|X_j, \tilde{P}_{j:E}, \pi, E) = 0$, $\tilde{P}_j$ identified from the regression of P_j on Z and $\pi(Z)$ vector of firms' entry probabilities identified from data

Robinson (1988): (β, α) identified (given σ)

σ identified from conditional moment on $\eta_{j:E}$, i.e. $\mathrm{E}[\eta_{j:E}|X_j, \tilde{P}_{j:E}, \pi, E] = 0$

Cost Parameters γ

Fix (j, E), conditioning supply equation on (Z, E) gives the partial linear regression model

$$\log\left[P_j - \frac{S_j}{\partial s_{j:C}(X_E, P_E, \xi_E; \theta)/\partial P_j}\right] = W_j'\gamma + \lambda^{\omega}_{j:E}(\pi) + \tilde{\omega}_{j:E}$$

with $E[\tilde{\omega}_{j:E}|W_j, \pi, E] = 0$, ξ_E and θ identified from above

Similar argument

Entry Costs Parameters κ

Entry equations involve the thresholds ξ^*, belief $F_{\xi_{E\setminus\{j\}}|\xi_j Z E \omega_E}(\cdot, \cdot|\xi_j^*, Z, E)$ and market structure probability $\Pr(E|\xi_j = \xi_j^*, Z)$ conditonal on (ξ_j^*, Z)

ξ_j^* identified as the lower bound of $F_{\xi_j|ZD_j}(\cdot|Z, 1)$ (identified from previously)

κ_j overidentified because of Z, or consider $\kappa_j(Z)$

Brand Effects μ

In partially linear regression model, the constant term is not identified.

$\rightarrow$ Robinson (1988)

$\rightarrow$ $\mu_j = \mathrm{E}(\xi_j)$ not identified

$F_{\xi_j}(\cdot)$ identified on $[\underline{\xi}_j^*, \overline{\xi}_j]$ only with $\underline{\xi}_j^* = \inf_z \xi^*(z)$

$\rightarrow$ Upper bound

$$\overline{\mu_j} = \underline{\xi}_j^* F_{\xi_j}(\underline{\xi}_j^*) + \int_{\underline{\xi}_j^*}^{+\infty} \xi_j dF_{\xi_j}(\xi_j)$$

Comment: If sufficient variation of $\xi_j^*(Z)$ in Z, μ_j can be identified under support condition $\sup_z \pi_j(z) = 1$.

Thus, μ_j is identified for those firms

Or for some Z all firms enter (no selection)

3. Estimation

ESTIMATION APPROACH

Combine IV with control function

Semiparametric (nonparametric first step for IV and seminonparametric correction terms)

Several possible estimators

DEMAND PARAMETERS $\theta = (\beta', \alpha, \sigma')'$

A Two-Step procedure: fix (j, E) (consider the subsample with the largest number of observations)

Step 1: a. Estimate π_j by a nonparametric regression of D_j on Z using full sample $\rightarrow \hat{\pi}_j$, $j = 1, \ldots, \overline{J}$

b. Estimate $\tilde{P}_{j:E}$ by a nonparametric regression of P_j on $(X_E, W_E, \hat{\pi})$ using subsample E $\rightarrow \hat{\tilde{P}}_{j:E}$

Remark: Oligopoly prices and entry probabilities depend on (X_E, W_E) and (X, W) only through the linear indices $(X_E'\beta, W_E'\gamma)$ and $(X'\beta, W'\gamma)$, respectively.

$\rightarrow$ Use of multiple-index models to estimate π and $\tilde{P}_{j:E}$ to reduce curse of dimensionality

Step 2: Consider $\lambda^{\xi}_{j:E}(\pi) = \sum_{k=1}^{K_\xi} \delta^{\xi}_{j:E,k} \phi^{\xi}_k(\pi)$ with $\phi^{\xi}_k(\cdot)$ basis functions

Minimum Distance Estimator based on

$$Q^d_{j:E}(\theta) = \sum \left[\psi_{j:C}(S;\sigma) - X'_j\beta + \alpha \hat{\hat{P}}_{j:E} - \sum_{k=1}^{K} \delta^{\xi}_{j:E,k} \phi^{\xi}_{j:E,k}(\hat{\pi}) \right]^2$$

by summing over all markets with $j : E$ (omitting index m)

$\to \hat{\theta}$

Cost Parameters γ

Fix (j, E)

Consider $\lambda^{\omega}_{j:E}(\pi) = \sum_{k=1}^{K_\omega} \delta^{\omega}_{j:E,k} \phi^{\omega}_k(\pi)$ with $\phi^{\omega}_k(\cdot)$ basis functions

Least Squares of $\log \left[P_j - \frac{S_j}{\partial s_{j:C}(X_E, P_E, \hat{\xi}_E; \hat{\theta})/\partial P_j} \right]$ on $[W_j, \phi^{\omega}_1(\hat{\pi}), ..., \phi^{\omega}_{K_\omega}(\hat{\pi})]$ on the subsample E to estimate γ

with $\hat{\xi}_E$, $\hat{\theta}$ and $\hat{\pi}$ from demand estimation

Pooled Estimation of Demand and Cost

Basic Idea: Pooling across firms and sets E to improve efficiency

Minimum Distance estimators based on

$$Q_M^d(\theta) \equiv \sum_{m=1}^{M} \sum_{j \in E_m} \left[\psi_{j:C_m}(S_m^*;\sigma) - X_{jm}'\beta + \alpha \hat{P}_{j:E_m} - \sum_{k=1}^{K_\xi} \delta_{j:E_m,k}^{\xi} \phi_{j:E_m k}^{\xi}(\hat{\pi}_m) \right]^2$$

$$Q_M^c(\hat{\theta},\gamma) \equiv \sum_{m=1}^{M} \sum_{j \in E_m} \left[\log\left(P_{jm} - \frac{S_{jm}}{\partial s_{j:C_m}(\mathbf{X}_{Em}, P_{Em}, \hat{\xi}_{Em}; \hat{\theta})/\partial P_j} \right) - W_{jm}'\gamma - \sum_{k=1}^{K_\omega} \delta_{j:E_m,k}^{\omega} \phi_{j:E_m k}^{\omega}(\hat{\pi}_m) \right]^2$$

where m is a market index and M total number of markets

Alternatively, estimate (θ, γ) simultaneously by considering $Q_M(\theta,\gamma) = Q_M^d(\theta) + Q_M^c(\theta,\gamma)$, where $\xi_{Em} = \{\psi_{j:C}(S_m;\sigma) - X_{jm}'\beta + \alpha P_{jm}; j \in E_m\}$

Entry Costs κ

Equation defining entry involves the thresholds ξ_j^*, firm's j, the belief $F_{\xi_{E\setminus\{j\}}\omega_E|\xi_j \mathbf{Z}\mathcal{E}}(\cdot,\cdot|\xi_j^*,\cdot,E)$ about competitors' qualities and entrants' cost shocks, and the probability $\Pr[\mathcal{E}=E|\xi_j=\xi_j^*,\mathbf{Z}=\cdot]$ of who will be firm j's opponents Instead we can view it as the expectation of $\Pi_{j:E}^* = \pi_{j:E}^*(X_E, W_E, \xi_j^*, \xi_{E\setminus\{j\}}, \omega_E; \theta, \gamma)$ on (ξ_j, Z) evaluated at (ξ_j^*, z) using the subsample corresponding to E.

1. Estimate $\hat{\xi}_j^*(Z) = \min \hat{\xi}_{j:E}(Z)$ over all observations of firm j given Z from demand estimation
2. Estimate $\mathrm{E}[\pi_{j:E}^*|\xi_j^*, Z, E]$ by a nonparametric regression of $\hat{\pi}_{j:E}^*$ on $(\hat{\xi}_j, Z)$ using subsample E evaluated at $\hat{\xi}_j^*(Z)$
3. Estimate $\Pr[E|\xi_j^*, Z]$ by a nonparametric regression of $1\!\!1(\mathcal{E} = E)$ on $(\hat{\xi}_j, Z)$ evaluated at $\hat{\xi}_j^*(Z)$

$\rightarrow$ Plug in estimated values/functions in entry equation gives $\hat{\kappa}_j$

Remarks:

(i) ξ_j^* is at boundary of nonparametric regressions in Step 2 (warning, boundary effects!)

(ii) To reduce curse of dimensionality, use $\lambda_{j:E}^{\pi}[\hat{\xi}_j^*, X_E'\beta, W_E'\gamma, \hat{\pi}_{-j}]$ in Step 2 and $\lambda_{j:E}^{\pi}[\hat{\xi}_j^*, \hat{\pi}_{-j}]$ in Step 3

(iii) $\hat{\kappa}_j(Z)$ can be averaged over markets to get $\hat{\kappa}_j$

Upper Bounds on Brand Effects $\overline{\mu}_j$

Noting that

$$F_{\xi_j}(\xi_j^*(Z)) = \Pr(D_j = 0|Z) = 1 - \pi_j(Z)$$

$$\int_{\xi_j^*(Z)}^{+\infty} \xi_j \partial F_{\xi_j}(\xi_j) = \mathrm{E}[\xi_j|Z, D_j = 1] \times \Pr[D_j = 1|Z] = \sum \mathrm{E}[\xi_j|Z, E] \times \Pr[E|Z]$$

where $\mathrm{E}[\xi_j|Z, E] = \lambda_{j:E}^{\xi}[\pi(Z)]$

$\pi_j(Z)$, $\xi_j^*(Z)$ and $\lambda_{j:E}^{\xi}[\pi(Z)]$ already estimated

It remains to estimate $\Pr[E|Z]$ from a nonparametric regression

$\rightarrow \hat{\overline{\mu}}_j$, $j = 1, \ldots, \overline{J}$

4. Monte Carlo Study

Set-Up

Logit Model with $\overline{J} = 2$

$X_j = 2, 2.5$ or 3 with equal probabilities, $W_j = -2.4 + 1.3X_j$

$(\xi_1, \xi_2)'$ normally distributed with mean $(\mu_1, \mu_2)' = (0.1, 0.0)'$, standard deviation $\sigma_\xi = 0.5$

Design 1: The ξs are independent, $\rho_\xi = 0$ and $\omega_j = -0.05$ or 0.35 with equal probabilities

Design 2: $\rho_\xi = 0.5$, $\omega_j = -0.05$ with probability $1/\{1 + \exp[(\xi_j - \mu_j)/\sigma_\xi]\}$ and 0.35 otherwise

$\beta = 1$, $\alpha = 0.5$, $\gamma = 0.5$, $\kappa_1 = \kappa_2 = 1.2$

Number of markets M=1,600 and K=500 replications

Data Generating Process

For each m, we need to know the market structure (no entry, monopoly 1, monopoly 2, duopoly), the price(s) P_j in case of entry and the corresponding market share S_j

With discrete (X, W), we need to compute 9 thresholds $\xi_j^*(Z)$ from entry equation

Comments:

(i) In Design 1 $f_{\xi_{-j}\omega|\xi_j}(\cdot, \cdot|\cdot) = f_{\xi_{-j}}(\cdot)f_{\omega_{-j}}(\cdot)f_{\omega_j}(\cdot)$ from independence

(ii) In Design 2 $f_{\xi_{-j}\omega|\xi_j}(\cdot, \cdot|\cdot) = f_{\omega_{-j}|\xi_{-j}}(\cdot|\cdot)f_{\omega_j|\xi_j}(\cdot|\cdot)f_{\xi_{-j}|\xi_j}(\cdot|\cdot)$

simplifying the computation of $\xi_1^*(\cdot), \xi_2^*(\cdot)$

(iii) Computation of profits $\pi_{mj}(\xi_j^*, \omega_j)$ and $\pi_{dj}(\xi_j^*, \xi_{-j}, \omega)$ which depend on monopoly and duopoly prices

Compare the draws (ξ_1, ξ_2) given Z to $(\xi_1^*(Z), \xi_2^*(Z))$

$\rightarrow$ Market Structure

$\rightarrow$ Prices P_j when entry

$\rightarrow$ Market Shares S_j when entry with $\psi_{j:C}(S^*; \sigma) = \log S_j/S_0$

Basis Functions

Main Idea: Develop a basis (i) reducing to Heckman correction when $(\xi_1, \xi_2, \omega_1, \omega_2)$ are iid normally distributed, and (ii) economizing on parameters

$\rightarrow$ Joint normality with truncation + terms for deviations for normality

$$\lambda^{\xi}_{j:\{1,2\}}(\pi) \;=\; \delta_0 + \delta_1 \frac{\phi(\zeta_j^*)[1-\Phi(\zeta_{-j}^*)]}{\pi_{12}} + \delta_2 \frac{\phi(\zeta_{-j}^*)[1-\Phi(\zeta_j^*)]}{\pi_{12}} + \frac{\phi(\zeta_1^*)\phi(\zeta_2^*)}{\pi_{12}}(\lambda_{00} + \lambda_{10}\zeta_1^* + \lambda_{01}\zeta_2^* + \ldots)$$

$$\lambda^{\xi}_{j:\{j\}}(\pi) \;=\; \delta_0 + \delta_1 \frac{\phi(\zeta_j^*)\Phi(\zeta_{-j}^*)}{\pi_{j0}} + \delta_2 \frac{\phi(\zeta_{-j}^*)[1-\Phi(\zeta_j^*)]}{\pi_{j0}} + \frac{\phi(\zeta_1^*)\phi(\zeta_2^*)}{\pi_{j0}}(\lambda_{00} + \lambda_{10}\zeta_1^* + \lambda_{01}\zeta_2^* + \ldots),$$

$\Phi(\cdot)(\phi(\cdot))$ cdf(pdf) of standard Normal, $\zeta_j^* \equiv \Phi^{-1}(1-\pi_j)$, $j = 1, 2$

$\pi_j = \Pr[D_j = 1|Z]$ (probability firm j entering)

$\pi_{12} = \Pr[D_1 = 1, D_2 = 1|Z]$ (probability duopoly), $\pi_{j0} = \Pr[D_j = 1, D_{-j} = 0|Z]$ (probability firm j monopoly)

Comments:

(i) $\{\delta_0, \delta_1, \delta_2, \lambda_{00}, \lambda_{10}, \lambda_{01}, \ldots\}$ need not be equal in the two correction terms

(ii) If (ξ_1, ξ_2) are iid Normal, $\lambda^{\xi}_{j:\{1,2\}}(\pi)$ and $\lambda^{\xi}_{j:\{j\}}(\pi)$ reduce to Heckman correction (δ_2, λ_{00}, $\ldots$vanish

$$\lambda^{\omega}_{j:\{1,2\}}(\pi) = \delta_0 + \delta_1 \frac{\phi(\zeta^*_j)[1-\Phi(\zeta^*_{-j})]}{\pi_{12}} + \delta_2 \frac{\phi(\zeta^*_{-j})[1-\Phi(\zeta^*_j)]}{\pi_{12}} + \frac{\phi(\zeta^*_1)\phi(\zeta^*_2)}{\pi_{12}}(\lambda_{00} + \lambda_{10}\zeta^*_1 + \lambda_{01}\zeta^*_2 + \dots)$$

$$\lambda^{\omega}_{j:\{j\}}(\pi) = \delta_0 + \delta_1 \frac{\phi(\zeta^*_j)\Phi(\zeta^*_{-j})}{\pi_{j0}} + \delta_2 \frac{\phi(\zeta^*_{-j})[1-\Phi(\zeta^*_j)]}{\pi_{j0}} + \frac{\phi(\zeta^*_1)\phi(\zeta^*_2)}{\pi_{j0}}(\lambda_{00} + \lambda_{10}\zeta^*_1 + \lambda_{01}\zeta^*_2 + \dots)$$

Comment (ii) applies when $(\xi_1, \xi_2, \omega_1, \omega_2)$ are iid Normal

Remark: The 4th term of the corrections constitute a complete basis, room for 3 constraints in every correction term!

Estimated Models

(M1) $\log S_j/S_0 = \beta X_j - \alpha \mathrm{E}[P_j|X] + \mu_j + \eta_j$

(M2) $\log S_j/S_0 = \beta X_j - \alpha \mathrm{E}[P_j|\mathbf{Z}, E] + \mu_{j:E} + \eta_j$

(M3) $\log S_j/S_0 = \beta X_j - \alpha \mathrm{E}[P_j|\mathbf{Z}, E] + \mu_j + \sigma_{\xi j} h(\pi_j) + \eta_{j:E}$

(M4) $\log S_j/S_0 = \beta X_j - \alpha \mathrm{E}[P_j|\mathbf{Z}, E] + \lambda^{\xi}_{j:E}(\pi) + \eta_{j:E}$

M1 estimated through 2SLS using X as instruments

M2 with a nonparametric first step with (X, W, E) as instruments and 4 constants $\mu_{j:E}$ as a 'naive' attempt to correct for endogenous choice set

M3 Heckman model $h(\pi_j) = \phi[\Phi^{-1}(1-\pi_j)]/\pi_j$ $(= \phi(\zeta_j^*)/[1-\Phi(\zeta_j^*)]$ with $\zeta_j^* = (\xi_j^* - \mu_j)/\sigma_\xi)$, nonparametric first step as in M2

M4 our model but room for 12 constraints, δ_0 and δ_1 equal, δ_2 opposite signs whether $\{j, 0\}$ or $\{1, 2\}$ (6 constraints), $\lambda_{00} = 0$ whether j is monopoly or duopoly (4 constraints), λ_{10} or λ_{10} opposite signs (2 constraints)

$\rightarrow$ M4a

$\rightarrow$ M4b

Estimate the πs by proportions given Z

Upper bounds on brand effects μ_j (in M4, M4a, M4b)

$$\hat{\bar{\mu}}_j \;=\; \min_z \left\{ \hat{\xi}_j^*(z)[1-\hat{\pi}_j(z)] + \hat{\lambda}^{\xi}_{j:\{j\}}[\hat{\pi}(z)]\hat{\pi}_{j0}(z) + \hat{\lambda}^{\xi}_{j:\{1,2\}}[\hat{\pi}(z)]\hat{\pi}_{12}(z) \right\}$$

with $\hat{\lambda}^{\xi}_{j:\{j\}}[\hat{\pi}(z)]$, $\hat{\lambda}^{\xi}_{j:\{1,2\}}[\hat{\pi}(z)]$ with parameter estimates $(\hat{\delta}_0, \hat{\delta}_1, \hat{\delta}_2, \hat{\lambda}_{00}, \hat{\lambda}_{10}, \hat{\lambda}_{01})$ from M4, 4a and 4b, $\hat{\xi}_j^*(z) = \min_{m:z_m=z}\{\hat{\xi}_{j:\{j\}}(z_m), \hat{\xi}_{j:\{1,2\}}(z_m)\}$ from $\hat{\xi}_{j:E}(z) = \log S_j/S_0 - \hat{\beta}x_j + \hat{\alpha}P_j$

Cost parameter γ

Same models as before with

$$\log\left[P_j - \frac{1}{\alpha(1-S_j)}\right] \;=\; W_j'\gamma + \lambda^{\omega}_{j:E}(\pi) + \tilde{\omega}_{j:E} \tag{1}$$

Replace α by $\hat{\alpha}$ and π by $\hat{\pi}$ obtained before

Entry costs κ

Regression of $\hat{\Pi}^*_{j:E} = \pi^*_{j:E}(X_E, W_E, \hat{\xi}^*_j, \hat{\xi}_{E\setminus\{j\}}, \hat{\omega}_E; \hat{\theta}, \hat{\gamma})$ and $1\!\!I(\mathcal{E}=E)$ on $(\hat{\xi}_{j:E}, Z)$

Because of boundary of ξ^*_j, one-sided uniform kernel $1\!\!I[0 \le x \le 1]$

$$\hat{\mathrm{E}}[\hat{\Pi}^*_{j:E}|\hat{\xi}_{j:E}=\hat{\xi}^*_j, Z=z, \mathcal{E}=E] \;=\; \frac{1}{M_{j:E}(z;h)} \sum_{m:E_m=E, z_m=z, \hat{\xi}_{j:E}\le\hat{\xi}^*_j+h} \hat{\Pi}^*_{j:E_m} \tag{2}$$

$$\widehat{\Pr}[\mathcal{E}=E|\hat{\xi}_{j:E}=\hat{\xi}_j^*, Z=z] = \frac{M_{j:E}(z;h)}{M_j(z;h)} \tag{3}$$

h bandwidth, $M_{j:E}(z;h)$ number of markets with competition E, characteristics z, estimated firm j quality with h of $\hat{\xi}_j^*$, similar for $M_j(z;h)$ without E

$$\rightarrow \quad \hat{\kappa}_j(z)=\frac{1}{M_j(z;h)} \sum_{m:j\in E_m, z_m=z, \hat{\xi}_{j:E}\leq\hat{\xi}_j^*+h} \hat{\Pi}^*_{j:E_m}$$

Monte Carlo Results

	π_1	π_2	π_{10}	π_{20}	π_{12}	$\text{corr}(\pi_1,\pi_2)$
Overall (Design 1)	0.545	0.510	0.290	0.255	0.254	-0.089
Overall (Design 2)	0.544	0.510	0.221	0.188	0.323	0.193

Positive correlation in Design 2 because of correlated ξs, also more monopolies

Design 1

		M1	M2	M3	M4	M4a	M4b
α	Mean	1.306	0.393	0.490*	0.403*	0.427*	0.457*
	STD	0.201	0.004	0.020	0.051	0.061	0.082
	RMSE	0.831	0.107	0.023	0.109	0.095	0.093
β	Mean	2.067	-0.156	0.978*	0.394*	0.775*	0.861*
	STD	0.304	0.062	0.043	0.313	0.195	0.205
	RMSE	1.110	1.157	0.049	0.682	0.298	0.248

Overestimates by BLP model, poor fit $R^2 < 0.1$

The naive correction does not work

Best results with Heckman (the ξs are independent!)

M4, 4a and 4b capture the true values

Design 2

		M1	M2	M3	M4	M4a	M4b
α	Mean	1.688	0.403	0.762	0.407	0.423*	0.425*
	STD	0.438	0.003	0.033	0.023	0.052	0.052
	RMSE	1.267	0.097	0.264	0.096	0.093	0.091
β	Mean	2.677	-0.091	1.388	0.392	0.720*	0.735*
	STD	0.674	0.071	0.062	0.257	0.202	0.197
	RMSE	1.808	1.094	0.393	0.660	0.346	0.330

Larger bias in M1 because selection is accentuated in Design 2

The Heckman correction is no longer appropriate because of the correlation

Important to impose restrictions

5. Extensions and Concluding Remarks

Forthcoming extensions: Endogenous characteristics X, random price coefficients,
incomplete information in both stages, multiproduct firms

Computationally friendly estimator to account for endogenous choice set

Testing for endogenous choice set (testing joint significance of correction terms parameters)

Empirical applications