Conduct

Chris Conlon

October 10, 2024

Grad IO

Conduct Testing in Industrial Organization

Foundational Empirical IO Question: How do we observe data on price and quantity and infer which model of firm behavior generated those outcomes?

Conduct Testing in Industrial Organization

Foundational Empirical IO Question: How do we observe data on price and quantity and infer which model of firm behavior generated those outcomes?

- Early work: Porter (1983), Bresnahan (1982,1987)
- Subsequent work defined the "menu" approach: Nevo (1998, 2001), Villas-Boas (2007)
- Recent revival of "internalization" parameters: Miller and Weinberg (2017), Crawford, Lee, Whinston, and Yurukoglu (2017), Pakes (2017)
- Parallel work by: Duarte, Magnolfi, Sølvsten, Sullivan (2022) which test is best (RV). Magnolfi, Quint, Sullivan, Waldfogel (2022) Should we test or estimate?
- Applications of our test: Starc and Wollman (2022), Scuderi (2022), others?

Is conduct testable? Berry and Haile (2014): yes.

Conduct Testing in Industrial Organization

- Absent additional restrictions, we cannot generally look at data on (P, Q) and decide whether or not collusion is taking place.
    - You say we started colluding at date t, I say we received a correlated shock to mc.
- We can make progress in two ways: (1) parametric restrictions on marginal costs; (2) exclusion restrictions on supply.
    - Most of the literature focuses on (1) by assuming something like: $\ln mc_{jt} = x_{jt}\gamma_1 + w_{jt}\gamma_2 + \omega_{jt}$.
    - In principle (2) is possible if we have instruments that shift demand for products but not supply. (These are much easier to come up with than "supply shifters").

A famous plot (Bresnahan 87)

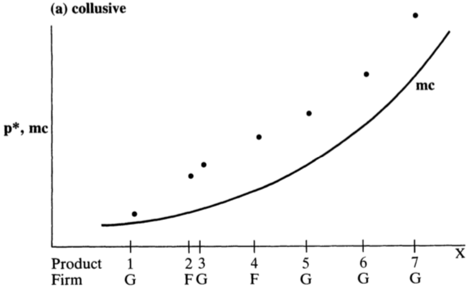


Figure 2(a)

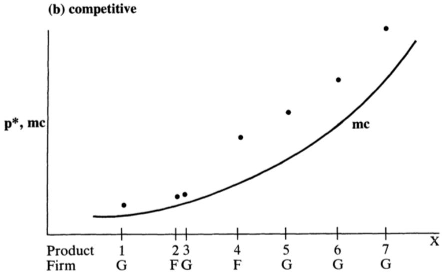
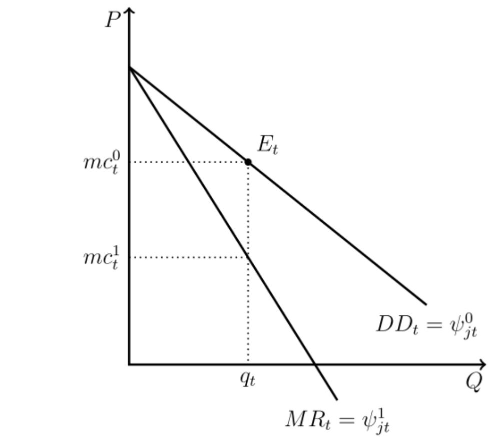


Figure 2(b)

Bresnahan (1980/1982) recognized this problem: we need "rotations of demand".

Conduct Testing in Pictures (Berry Haile 2014)



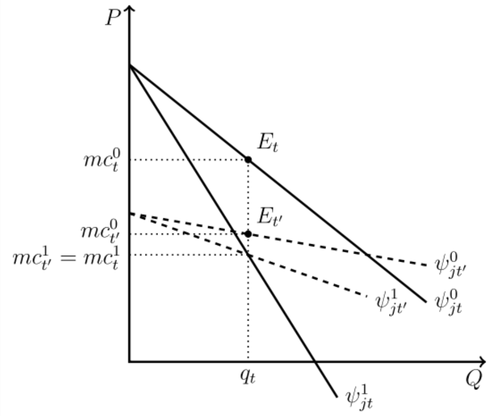


Figure 2(ab) from Berry and Haile (2014), Example 1.

Setup: Notation and Utility

We begin with a relatively standard BLP-style differentiated products setup.

- Markets t
- Products j
- Data $\chi_t = \{(\mathrm{x}_{jt}, \mathrm{v}_{jt}, \mathrm{w}_{jt})$ for all $j \in \mathcal{J}_t\}$.
- Market Shares $\mathcal{S}_t = [s_{1t}, \ldots, s_{Jt}, s_{0t}]$.
- Prices $\mathbf{p}_t = [p_{1t}, \ldots, p_{Jt}]$.
- Consumers i with demographics y_{it} (income, presence of kids)

Testing Conduct: Setup

We generalize the $\mathcal{H}(\kappa)$ and derive multi-product Bertrand FOCs:

$$\arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p}) = \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot q_j(\mathbf{p}) + \kappa_{fg} \sum_{j \in \mathcal{J}_g} (p_j - c_j) \cdot q_j(\mathbf{p})$$

$$\rightarrow 0 = q_j(\mathbf{p}) + \sum_{k \in (\mathcal{J}_f, \mathcal{J}_g)} \kappa_{fg} \cdot (p_k - c_k) \frac{\partial q_k}{\partial p_j}(\mathbf{p})$$

- Instead of 0's and 1's we now have $\kappa_{fg} \in [0, 1]$ representing how much firm f cares about the profits of g.
    - If f and g merge (or fully coordinated) then $\kappa_{fg} = 1$
    - Often in the real world firms cannot reach fully collusive profits and $\kappa_{fg} \in (0, 1)$.
    - Evidence that $\kappa_{fg} > 0$ is not necessarily evidence of malfeasance, just a deviation from static Bertrand pricing

Testing Conduct: Setup

- Recall the Δ matrix which we can write as $\Delta = \tilde{\Delta} \odot \mathcal{H}(\kappa)$, where $\odot$ is the element-wise or Hadamard product of two matrices.
    - $\tilde{\Delta}$ is the matrix of demand derivatives with $\Delta(j,k) = \frac{\partial q_j}{\partial p_k}$ for all elements.
    - $\mathcal{H}(\kappa) = \kappa_{fg}$ for products owned by (f,g) where $\kappa_{ff} = 1$ always.
- Mergers are about changing 0's to 1's in the $\mathcal{H}(\kappa)$ matrix.
- Matrix form of FOC: $q(\mathbf{p}) = \Delta(\mathbf{p},\kappa) \cdot (\mathbf{p} - \mathbf{mc})$
- $\mathbf{mc} = \mathbf{p} - \underbrace{\Delta(\mathbf{p},\theta_2,\kappa)^{-1} s(\mathbf{p})}_{\eta(\mathbf{p},\mathbf{s},\theta_2,\kappa)}$ where η_{jt} is the markup.

Reasons for Deviations from Static Bertrand

Biased estimates of own and cross price derivatives: For anything to work, you have correct estimates of $\tilde{\Delta}$. My prior is most papers underestimate diversion ratios for close substitutes.

Vertical Relationships: Who sets supermarket prices? Just the retailer? Just the manufacturer? Some combination of both? Retailers tend to soften downstream price competition.

Faulty Timing Assumptions: Bertrand is a simultaneous move pricing game. Lots of alternatives (Stackelberg leader-follower, Edgeworth cycles, etc.).

Dynamics and Dynamic Pricing: Forward looking firms or consumers might not set static Nash prices. [e.g. Temporary Sales, Switching Costs, Network Effects, etc.]

Unmodeled Supergame: Maybe firms are legally tacitly colluding, higher prices might be about what firms believe will happen in a price war.

Simultaneous Problem

Assume additivity, and write in terms of structural errors:

$$\delta_{jt}(\mathcal{S}_t, \widetilde{\theta}_2) + \alpha p_{jt} = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}, \theta_1) + \xi_{jt}$$
$$f\left(p_{jt} - \eta_{jt}(\mathbf{p}, \mathbf{s}, \theta_2, \kappa)\right) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3) + \omega_{jt}$$

- To simplify slides we let $f(x) = x$ (often $f(x) = \log(x)$) but we can put that in $h_s(\cdot)$.
- $h(\cdot)$ are often just linear relationships like: $\theta_1[\mathrm{x}_{jt}, \mathrm{v}_{jt}]$.
- Endogeneity Problem: p_{jt} and η_{jt} are functions of $(\boldsymbol{\xi}, \boldsymbol{\omega})$.
- (θ_2, κ) parameters that determine markups

Approach #1: Demand Side

1. Estimate θ_2 from demand alone.

$$\delta_{jt}(\mathcal{S}_t, \widetilde{\theta}_2) + \alpha p_{jt} = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}, \theta_1) + \xi_{jt}$$
$$E[\xi_{jt} | \mathrm{x}_t, \mathrm{v}_t, \mathrm{w}_t] = 0$$

2. Recover marginal costs $\widehat{\mathbf{mc}} = \mathbf{p} + \boldsymbol{\eta}$

$$\boldsymbol{\eta}(\mathbf{p}, \mathbf{s}, \theta_2, \kappa) \equiv \left(\mathcal{H}(\kappa) .* \tilde{\Delta}(\mathbf{p}, \theta_2)\right)^{-1} q(\mathbf{p})$$

Challenges:

- Given $[\mathbf{q}, \mathbf{p}, \tilde{\Delta}, \mathcal{H}(\kappa)]$ I can always produce a vector of marginal costs $\mathbf{mc}$ that rationalizes what we observe. [ie: J equations J unknowns].
- Nonparametrically we cannot identify κ without more restrictions (!).

What do people do?

Maybe some vectors of $\mathbf{mc}$ look less "reasonable" than others.

- Marginal costs ≤ 0 seem problematic. [Might just be that your estimates for demand are too inelastic...]
- or I have a parametric model of MC in mind.

$$f\left(p_{jt} - \eta_{jt}(\mathbf{p}, \mathbf{s}, \theta_2, \kappa)\right) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3) + \omega_{jt}$$
$$E[\omega_{jt} | \mathrm{x}_t, \mathrm{w}_t, \mathrm{v}_t] = 0$$

- Can test that model with GMM objective of mc_{jt} on regressors.
- Maybe marginal costs cannot deviate too much within product from period to period. (We can write these as moment restrictions too).

Approach #2: Simultaneous Supply and Demand

Estimate θ_2 using both supply and demand. The fit of my supply side will also inform my demand parameters, particularly α the price coefficient. [BLP 95 used this for additional power with lots of random coefficients and potentially weak instruments].

$$\delta_{jt}(\mathcal{S}_t, \widetilde{\theta}_2) + \alpha p_{jt} = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}, \theta_1) + \xi_{jt}$$
$$f\left(p_{jt} - \eta_{jt}(\mathbf{p}, \mathbf{s}, \theta_2, \kappa)\right) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3) + \omega_{jt}$$

Challenges:

- Should I try to estimate κ? or just compare objective values at $\kappa_{fg} \in \{0, 1\}$?
- Am I testing conduct? Or am I testing the functional form for my supply model?
- Will a missing IV/restriction change whether or not I believe firms are colluding?

Estimation vs. Menus

There are two ways to think about conduct:

1. Using moment conditions to estimate $\widehat{\kappa}$ or $\mathcal{H}(\kappa)$ directly.
    - Often with a small number of parameters (ie: $\kappa_{fg} = 0$ except for firms I know are in a cartel).
    - Can be challenging to tell similar values of κ_{fg} apart (under-powered).
2. "Menu Approach"
    - Nevo (Economics Letters 1998)
    - Bresnahan (1987)
    - Compare some goodness-of-fit critieria across assumed values of κ (Bertrand vs. Collusion)
3. Rejecting (or failing to reject) a single model.

Testing a single model of κ

Put the η_{jt} on the RHS and test whether $\lambda = 1$:

$$p_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3) + \lambda \cdot \eta_{jt}\left(\mathbf{p}, \mathbf{s}, \theta_2, \kappa\right) + \omega_{jt} \text{ with } E[\omega_{jt} | \mathrm{x}_t, \mathrm{w}_t, \mathrm{z}_t] = 0s$$

- We are basically running 2SLS with IV for the endogenous η_{jt}
- "Informal" test of Villas Boas (2007): $\mathbb{E}[\omega_{jt} | \mathrm{x}_{jt}, \mathrm{w}_{jt}, \eta_{jt}] = 0$.
  - Considers different forms of $f(\cdot)$: linear, exponential, logarithmic.
  - Not sure the published paper includes these results (?) WP does?
- Pakes (2017) uses Wollman (2018) data and BLP IV $\mathbb{E}[\omega_{jt} | x_{jt}, w_{jt}, f(x_{-j})] = 0$.
- $\lambda \neq 0$ is hard to interpret.

Pakes (2018)/Wollman (2017) Regressions: Heavy Trucks

Table 1: Wollman & Pricing Equilibrium.

Taken from Pakes, 2017, *Journal of Industrial Economics.*

	Price	(S.E.)	Price	(S.E.)
Gross Weight	.36	(0.01)	.36	(.003)
Cab-over	.13	(0.01)	.13	(0.01)
Compact front	-.19	(0.04)	0.21	(0.03)
long cab	-.01	(0.04)	0.03	(0.03)
Wage	.08	(.003)	0.08	(.003)
$\widehat{Markup}$	.92	(0.31)	1.12	(0.22)
Time dummies?	No	n.r.	Yes	n.r.
R^2	0.86	n.r.	0.94	n.r.

Note. There are 1,777 observations; 16 firms over the period 1992-2012. S.E.=Standard error.

Single Model Regressions

These are somewhat reassuring:

- $\lambda \approx 1$ for multiproduct-oligopoly
- Fit is pretty good $R^2 > 0.8$ and $R^2 > 0.5$ for within vehicle regressions (not shown).
- As a behavioral model, multiproduct demand estimation seems successful.
- But, do we know that an alternative $\mathcal{H}(\kappa)$ would have a $\lambda \neq 1$ or a lower R^2, and if so how low before we can "reject" the model?

Goodness of Fit Tests

Another idea (Bonnet and Dubois, Rand 2010) runs the following regression:

$$\log\left(p_{jt} - \eta_{jt}(\mathbf{p}, \mathbf{s}, \widehat{\theta}_2, \kappa)\right) = h_s(x_{jt}, w_{jt}, \theta_3) + \omega_{jt}$$

- Run a regression for each κ and obtain $Q(\kappa) = \sum_{jt} \widehat{\omega}_{jt}^2$
- Employ the non nested test of Rivers and Vuong (2002). Why?
- Working out the distribution of $Q(\kappa_1) - Q(\kappa_2) = T(\kappa_1, \kappa_2)$ is the hard part.
- Also this is OLS (or NLLS) and there are no instruments or exclusion restrictions for the supply side. Presumably we could add some and do GMM? (I think this is the "formal" test of Villas Boas (ReStud 2007)).

Recap

So far three approaches to exploit $E[\omega_{jt}|x_t, w_t, z_t] = 0$

1. Put the markup on RHS and instrument for it to test $\lambda = 1$ (Wald)

$$p_{jt} = h_s(x_{jt}, w_{jt}, \theta_3) + \lambda \cdot \eta_{jt}\left(\mathbf{p}, \mathbf{s}, \widehat{\theta_2}, \kappa\right) + \omega_{jt}$$

2. Put the markup on LHS assuming $\lambda = 1$ and test goodness of fit of supply equation (Anderson Rubin)

$$p_{jt} - \eta_{jt}\left(\mathbf{p}, \mathbf{s}, \widehat{\theta_2}, \kappa\right) = h_s(x_{jt}, w_{jt}, \theta_3) + \omega_{jt}$$

3. Estimate supply and demand simultaneously $[\theta_1, \theta_2, \theta_3]$ and compare goodness of fit for different κ. (Likelihood Ratio)

Backus Conlon Sinkinson (2022)

Basic Setup

We start with marginal revenue and marginal cost (unobserved ω, observed $h(\cdot)$)

$$\psi_{jt}^{m} = mc_{jt}$$
$$p_{jt} - \eta_{jt}^{m} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \omega_{jt}^{m}$$

- Let's be vague/flexible with $h_s(\cdot)$ for now, but I don't know the production function.
- Assume: Demand and hence η_{jt}^{m} are known (given conduct).
- Idea (η^A, η^B) are monopoly/perfect competition or Cournot/Bertrand.

The Question

Two competing markups $(\eta_{jt}^A, \eta_{jt}^B)$: which fits the data better?
(both may be misspecified)

$$p_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \tau\, \eta_{jt}^A + (1-\tau)\, \eta_{jt}^B + \omega_{jt}$$

Model is defined by a conditional moment restriction $\mathbb{E}[\omega_{jt} | z_{jt}^s] = 0$

- $H_0 : \tau = 1$ vs $H_a : \tau = 0$
- This is a model selection problem or a non nested testing problem.
  - We might want to compare more than two alternatives (too bad).
- Obvious endogeneity problem with η_{jt}!

Testing Environment

Compare violations of unconditional moments under $(\eta^A_{jt}, \eta^B_{jt})$ and $A(z^s_{jt})$:

$$p_{jt} - \eta^A_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \omega^A_{jt}$$
$$p_{jt} - \eta^B_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \omega^B_{jt}$$

Testing Environment

Compare violations of unconditional moments under $(\eta^A_{jt}, \eta^B_{jt})$ and $A(z^s_{jt})$:

$$p_{jt} - \eta^A_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \omega^A_{jt}$$
$$p_{jt} - \eta^B_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}) + \omega^B_{jt}$$

Which gives us

$$g_A = \frac{1}{N}\sum_{jt} \omega^A_{jt}\, A(z^s_{jt}), \quad g_B = \frac{1}{N}\sum_{jt} \omega^B_{jt}\, A(z^s_{jt})$$
$$Q_m = g'_m\, W_m\, g_m$$

Now consider a Rivers Vuong (2002) type test $T_{RV} = \sqrt{n}\left(\frac{Q_A - Q_B}{\sigma_{Q_A - Q_B}}\right) \sim N(0,1)$
$H_0 : Q_A - Q_B = 0$ vs. $H_A : Q_A > Q_B$ or $Q_A < Q_B$.
Getting the SD of the difference is hard $\rightarrow$ bootstrap

Comparison to Literature

- Bresnahan (1987): Did LR test to determine collusion vs. competition in 1955 automobile price war
  - No IV, errors were measurement in P, Q.
- Bonnet and Dubois (2010): RV test
  - But no IV – maximum likelihood with normally distirbuted ω_{jt}'s.
- Villas Boas (2007): Cox test to determine double marginalization or not in yogurt
  - GMM objective, unclear what if any IV are used.
  - Need to "know" the true model.
- Duarte, Magnolfi, Solvsten, Sullivan (2022): RV beats Cox pretty badly in Monte Carlo.

Main Results: These are $N(0,1)$

	Others' Cost	Demographics	BLP Inst.	Dmd. Opt. Inst.
Own Profit Max vs.	Panel 1: $A(\mathbf{z}_t) = \mathbf{z}_t$, linear $h_s(\cdot)$			
Common Ownership	-4.3410	-1.1966	0.5047	-1.2552
Double Marginalization	2.1922	1.0055	-0.0412	7.0897
Double Marginalization + CO	-0.8262	0.6892	0.1428	6.9320
Perfect Competition	3.2995	0.5194	0.7355	3.7223
Monopolist	-2.2264	-1.0528	-0.4525	-0.9202
Own Profit Max vs.	Panel 2: $A(\mathbf{z}_t) = \mathbb{E}[\Delta\eta^{12}\vert\mathbf{z_t}]$, linear $h_s(\cdot)$ and $g(\cdot)$			
Common Ownership	-2.3044	-0.5105	-0.0384	-1.6133
Double Marginalization	0.8644	0.4421	-0.5311	3.3367
Double Marginalization + CO	-0.9382	-0.2389	-0.3684	-0.0045
Perfect Competition	0.7164	0.6135	-0.1080	-0.3151
Monopolist	-0.8577	-0.4002	-0.3868	-1.2339
Own Profit Max vs.	Panel 3: $A(\mathbf{z}_t) = \mathbb{E}[\Delta\eta^{12}\vert\mathbf{z_t}]$, random forest $h_s(\cdot)$ and $g(\cdot)$			
Common Ownership	-3.3777	-3.2509	-3.7130	-4.0256
Double Marginalization	-5.9699	-9.9547	-6.5789	-7.8269
Double Marginalization + CO	-5.9264	-6.1550	-6.5231	-7.4760
Perfect Competition	-4.0468	-6.1901	-5.1494	-6.3484
Monopolist	-3.4972	-4.0070	-3.4358	-3.7495

An Internalization Parameter

Let κ represent the weight a firm places on competitors and τ the internalization of those weights.

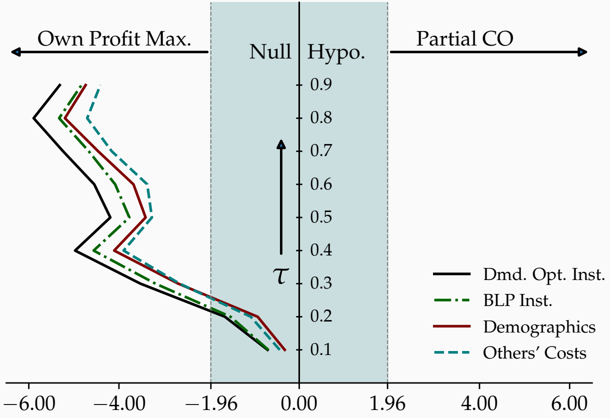
$$arg \max_{p_j : j \in \mathcal{J}_f} \sum_{j \in \mathcal{J}_f} (p_j - mc_j) \cdot s_j(\mathbf{p}) + \sum_{g \neq f} \tau \kappa_{fg} \sum_{j \in \mathcal{J}_g} (p_k - mc_k) \cdot s_k(\mathbf{p})$$

Now,

- $\tau = 0$ implies own-profit maximization
- $\tau = 1$ implies common ownership pricing
- τ in between is..? Agency?

We test $\tau \in (0.1, \ldots, 0.9)$ against own-profit maximization.

Internalization Parameter Results



Setup: Challenges

The true model for markups (conduct) will satisfy the CMR: $\mathbb{E}[\omega_{jt}|z^s_{jt}] = 0$

$$p_{jt} - \eta^{(m)}_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Goal is test two competing markups $\eta^{(A)}_{jt}, \eta^{(B)}_{jt}$, but there are challenges:

Setup: Challenges

The true model for markups (conduct) will satisfy the CMR: $\mathbb{E}[\omega_{jt}|z^s_{jt}] = 0$

$$p_{jt} - \eta^{(m)}_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Goal is test two competing markups $\eta^{(A)}_{jt}, \eta^{(B)}_{jt}$, but there are challenges:

1. Test will depend on how we choose unconditional moment restrictions $\mathbb{E}[\omega_{jt} \cdot A(z^s_{jt})] = 0$

Setup: Challenges

The true model for markups (conduct) will satisfy the CMR: $\mathbb{E}[\omega_{jt}|z^s_{jt}] = 0$

$$p_{jt} - \eta^{(m)}_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Goal is test two competing markups $\eta^{(A)}_{jt}, \eta^{(B)}_{jt}$, but there are challenges:

1. Test will depend on how we choose unconditional moment restrictions $\mathbb{E}[\omega_{jt} \cdot A(z^s_{jt})] = 0$
2. Test may depend on how we specify $h_s(\cdot)$
   - All tests are basically joint tests of the specification for observed marginal costs and the exclusion restriction.
   - Villas Boas (2007) tries log, linear, exponential in $x\beta$

Setup: Challenges

The true model for markups (conduct) will satisfy the CMR: $\mathbb{E}[\omega_{jt}|z^s_{jt}] = 0$

$$p_{jt} - \eta^{(m)}_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Goal is test two competing markups $\eta^{(A)}_{jt}, \eta^{(B)}_{jt}$, but there are challenges:

1. Test will depend on how we choose unconditional moment restrictions $\mathbb{E}[\omega_{jt} \cdot A(z^s_{jt})] = 0$
2. Test may depend on how we specify $h_s(\cdot)$
    - All tests are basically joint tests of the specification for observed marginal costs and the exclusion restriction.
    - Villas Boas (2007) tries log, linear, exponential in $x\beta$
3. Choice of $\eta^{(m)}_{jt}$ will affect our choice of weighting matrix and thus the test. (Hall Pelletier (2011))

A Brief Aside: Chamberlain (1987) in a Slide

What contains as much information as the CMR $\mathbb{E}[\omega|z^s_{jt}]$ and moments of the form $\mathbb{E}[\omega_{jt} \cdot A(z^s_{jt})]$.

- For linear models $A(z^s_{jt}) = z^s_{jt}$ is generally without loss.
- For nonlinear models, Chamberlain (1987) shows that the efficient estimator uses

$$A(z^s_{jt}) = \mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \theta}|z^s_{jt}\right]$$

- That is not too helpful (its a function of the unknown θ).
- Much of the follow-up work has been about feasible approximations to this "optimal instrument" (e.g., Newey 1990)

For us a similar concern arises, but it is about power to distinguish conduct models rather than efficiency of estimation.

Our Idea: Motivation #1 (Optimal IV)

The model is given by

$$p_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3) + \tau \cdot \eta^A_{jt} + (1-\tau) \cdot \eta^B_{jt} + \omega^m_{jt}$$

$$\text{where } H_0 : \tau = 1 \text{ and } H_a : \tau = 0$$

- The optimal IV in the Chamberlain (1987) sense is given by $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \tau} | z_t\right] = \mathbb{E}\left[\eta^A_{jt} - \eta^B_{jt} | z_t\right]$.
- In words: The IV need to predict the difference in markups (beyond observed $h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, \theta_3)$).

Our Idea: Motivation #2 (Misspecification)

Index the true model by 0. Then,

$$p_{jt} - \eta^0_{jt} = h_s(x_{jt}, w_{jt}) + \omega^0_{jt}.$$

To motivate a useful test, we ask what happens when we estimate supply with the wrong conduct model (1):

Our Idea: Motivation #2 (Misspecification)

Index the true model by 0. Then,

$$p_{jt} - \eta^0_{jt} = h_s(x_{jt}, w_{jt}) + \omega^0_{jt}.$$

To motivate a useful test, we ask what happens when we estimate supply with the wrong conduct model (1):

$$p_{jt} - \eta^1_{jt} = h_s(x_{jt}, w_{jt}) + \underbrace{\underbrace{\eta^0_{jt} - \eta^1_{jt}}_{\equiv \Delta\eta^{0,1}_{jt}} + \omega^0_{jt}}_{\omega^1_{jt}}.$$

- Misspecifying conduct introduces an omitted variable: the difference in markups.
- Our test is premised on detection of this omitted variable.

Our Innovation: How does this help?

The model is given by

$$p_{jt} - \eta_{jt}^m = h_s(\cdot) + \omega_{jt}^m, \text{ and } \mathbb{E}[\omega_{jt}^{(m)} \cdot A(z_t)] = 0.$$

We suggest $A(z_t) = \mathbb{E}[\eta_{jt}^1 - \eta_{jt}^2 | z_t]$; several advantages:

Our Innovation: How does this help?

The model is given by

$$p_{jt} - \eta_{jt}^{m} = h_s(\cdot) + \omega_{jt}^{m}, \text{ and } \mathbb{E}[\omega_{jt}^{(m)} \cdot A(z_t)] = 0.$$

We suggest $A(z_t) = \mathbb{E}[\eta_{jt}^{1} - \eta_{jt}^{2} | z_t]$; several advantages:

- Reduces potentially many moments ($\mathbb{E}[\omega_{jt}' z_t] = 0$) to a single, scalar moment. No need for a weighting matrix, or associated problems.
- Testing is reduced to two prediction exercises: $\mathbb{E}[\eta_{jt}^{1} - \eta_{jt}^{2} | z_t]$ and $\widehat{\omega}_{jt}^{(m)}$.
- Show in the paper that this leads to the most powerful test (maximizes distance between two GMM objective functions conditional on weight matrix).
- Downside: Our choice of instrument is model specific! UMP is not going to happen.

Possible Exclusion Restrictions

We are looking for variables which affect demand but not supply:

$$\sigma_j^{-1}(\mathcal{S}_t, \mathbf{p_t}, \mathrm{y_t}, \mathrm{x}_t, \mathrm{v}_t, \widetilde{\theta}_2) = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}; \theta_1) - \alpha\, p_{jt} + \lambda \log(\mathsf{ad}_{jt}) + \xi_{jt}$$

$$p_{jt} - \eta_{jt}(\mathcal{S}_t, \mathbf{p_t}; \theta_2, \mathcal{H}_t(\kappa)) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Things we use:

- Obvious choice: v_{jt} (things like product recalls are relatively weak)
- Demographics (enter nonlinearly): y_t (chain-level income works well)
- Characteristics of other goods: $f(\mathrm{x}_{-j,t})$ (BLP instruments).
- Characteristics of other goods: $\mathrm{w}_{-j,t}$ (commodity price of oats for Rice Krispies)

Possible Exclusion Restrictions

We are looking for variables which affect demand but not supply:

$$\sigma_j^{-1}(\mathcal{S}_t, \mathbf{p_t}, \mathrm{y_t}, \mathrm{x}_t, \mathrm{v}_t, \widetilde{\theta}_2) = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}; \theta_1) - \alpha\, p_{jt} + \lambda \log(\mathsf{ad}_{jt}) + \xi_{jt}$$

$$p_{jt} - \eta_{jt}(\mathcal{S}_t, \mathbf{p_t}; \theta_2, \mathcal{H}_t(\kappa)) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt}$$

Things we use:

- Obvious choice: v_{jt} (things like product recalls are relatively weak)
- Demographics (enter nonlinearly): y_t (chain-level income works well)
- Characteristics of other goods: $f(\mathrm{x}_{-j,t})$ (BLP instruments).
- Characteristics of other goods: $\mathrm{w}_{-j,t}$ (commodity price of oats for Rice Krispies)

Things we don't use:

- Unobserved demand shocks ξ_{jt} (see MacKay Miller 2020 for $Cov(\xi_j, \omega_j) = 0$).
- Observable κ conduct shifters (financial mergers/events, see Miller Weinberg (2018))

Algorithm

(1) Split the sample by markets t into 70% *test* and 30% *train*.

(2) On the *training sample*:

  (a) Approximate the optimal instruments $a(z^s_{jt}) = \mathbb{E}[\Delta\eta^{(1,2)}_{jt} \mid z^s_{jt}]$ as the fitted values from:

$$\Delta\eta^{1,2}_{jt} = g(z^s_{jt}) + \zeta_{jt}.$$

  (b) Estimate the marginal cost function, under models 1 and 2 to obtain residuals $\widehat{\omega^1_{jt}}$ and $\widehat{\omega^2_{jt}}$:

$$p_{jt} - \eta^m_{jt} = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega^m_{jt}.$$

(3) On the *test sample*:

  (a) For each candidate model, compute the value of the scalar moment:[1]

$$Q(\eta^m) = \left(\sum_{j,t} \hat{\omega}^m_{jt} \cdot \hat{g}(\mathbf{z_t})\right)^2.$$

  (b) Repeat the previous step on bootstrapped samples and estimate $\hat{\sigma}/\sqrt{n}$ the standard error of the difference $\tilde{Q}(\eta^1) - \tilde{Q}(\eta^2)$.

  (c) Compute the test statistic

$$T = \frac{\sqrt{n}\,(Q(\eta^1) - Q(\eta^2))}{\widehat{\sigma}} \sim \mathcal{N}(0,1).$$

Note: Steps 2(a) and 2(b) can be done in any order via non-parametric regression.

Limitations

Not everything is testable:

- If $\Delta\eta_{jt}$ cannot be explained by z^s_{jt} beyond contents of $(\mathrm{x}_j, \mathrm{w}_j)$ we have nothing
- Flexible demand models are required to generate cross sectional variation in markups
- Beware of “accidental” exclusion restrictions.

Alternate Justifications

1. Model Misspecification: if one of the two models is correct, $\eta^1_{jt} - \eta^2_{jt}$ exactly corresponds to the misspecification error when using the other. ⏩ detail.
2. Difference in Test Statistics: We can also show it maximizes the difference in GMM objectives $(Q_1 - Q_2)$. (Which is almost power for fixed σ). Also we need that either $h_s(\cdot)$ doesn't depend on $\eta^{(m)}$ or that it is linear so we can residualize on (x, w).