Lecture 5: Pricing Equilibria and Mergers

Chris Conlon

October 1, 2019

Grad IO

Antitrust and Mergers

What is Antitrust?

- Current Debate: should we maximize efficiency/social surplus or should we focus on consumer surplus.
  - US antitrust law focuses primarily on harm to consumers.
  - EC tends to also worry about harm to competing firms.
- We know about DWL from market power from undergrad economics. However, without profits, why would firms innovate or perform R&D?
  - Law understands this and awards temporarily monopolies via patents.
- Today, I am going to focus mostly on horizontal mergers among competitors.
  - Most of this is known as unilateral effects (which is a terrible name).
  - Also worry about coordinated effects which mean the nature of equilibrium changes.

Antitrust Legislation : Sherman Act (1890)

Section 1 "Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal" (Violation involves an agreement).

Section 2 "Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony".

Three *per se* violations

- (1) price fixing (2) horizontal market division (3) refusals to deal.
- Other violations are *rule of reason*.

Antitrust Legislation : Clayton Act (1914)

Section 2 Prohibits some forms of price discrimination, but only when it lessens competition.

Section 3 Prohibits sales based on the condition that the buyer not buy from your competitor (includes tying and exclusive dealing), but only when effect may be to substantially lessen competition.

Section 7 Prohibits mergers where the effect of such acquisition may be substantially to lessen competition, or tend to create a monopoly in any line of commerce.

Section 8 Prevents a person from being a director of multiple competing firms.

Antitrust Legislation : Hart-Scott-Rodino Act (1976)

- Required pre-notification and registration of large mergers
  - Transaction: \$78.2 million
  - Size of Person: \$156.3 M with target of \$15.6 M or total transaction of \$312.6M
  - These are "inflation adjusted" each year.
- Initial review period is 30 days after which DOJ/FTC can request additional information or allow merger to proceed.
- Second review usually involves detailed information about price-cost margins, market shares, etc. (Usually more info available than to academic researchers).
- Can request information company would reasonably have (customer surveys, etc.).
- After second review can ask for injunctive relief or remedies which merging parties can oppose in court.

DOJ/FTC Horizontal Merger Guidelines

- DOJ/FTC describe markets as:
  - Highly Concentrated: $HHI \geq 2500$.
  - Moderately Concentrated: $HHI \in [1500, 2500]$. $\Delta HHI \geq 250$ merits scrutiny.
  - Un-Concentrated: $HHI \leq 1500$.
- Also consider unilateral effects/UPP and coordinated effects.
- Three steps:
  1. Market Definition
  2. Measure Concentration/Initial Screening
  3. Merger Simulation

Step 1: Market Definition

SSNIP

- Small but significant and non-transitory increase in price (SSNIP): smallest relevant market where a hypothetical monopolist could impose a 5% price increase. (For at least one year).
- Under linear demand this amounts to a price cost margin and an elasticity (sometimes the critical elasticity).

Tricky Examples:

- Whole Foods vs. Wild Oats
- Cellophane Fallacy.

Step 2: Concentration/Screening

- After we define the relevant market, compute the relevant HHI or UPP.
- There can be both geographic and product market issues in the relevant market.
- Some markets may be highly concentrated and others may not be.
- Can ask for divestitures as part of a remedy if there are a few problematic markets in an otherwise uncontroversial merger.

Step 3: Merger Simulation

- Simulate the price effects of the merger
- Take into account likely cost synergies (sometimes there are none).
- Estimate post-merger prices and welfare.

This is what we will talk about next.

Merger Simulation

Merger Simulation: Two Options

- Partial Merger Simulation
  - Simulate a new price for p_j after acquiring good k holding the prices of all other goods (p_k, p_{-j}) fixed.
  - Repeat for p_k and all other products involved in the merger.
  - Compare price increases to synergies or cost savings.
- Full Merger Simulation
  - Write down the full system of post-merger FOC.
  - Adjust post-merger marginal costs for potential synergies.
  - Solve for all prices at the new (post-merger) equilibrium (p_j, p_k, p_{-j}).

Differentiated Products Bertrand

Recall the multi-product Bertrand FOCs:

$$\begin{aligned} \arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p}) &= \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot q_j(\mathbf{p}) \\ \rightarrow 0 &= q_j(\mathbf{p}) + \sum_{k \in \mathcal{J}_f} (p_k - c_k) \frac{\partial q_k}{\partial p_j}(\mathbf{p}) \end{aligned}$$

It is helpful to define the matrix Ω with entries:

$$\Omega_{(j,k)}(\mathbf{p}) = \left\{ \begin{array}{cl} -\frac{\partial q_j}{\partial p_k}(\mathbf{p}) & \text{for } (j,k) \in \mathcal{J}_f \\ 0 & \text{for } (j,k) \notin \mathcal{J}_f \end{array} \right\}$$

We can re-write the FOC in matrix form:

$$q(\mathbf{p}) = \Omega(\mathbf{p}) \cdot (\mathbf{p} - \mathbf{mc})$$

Merger Simulation

What does a merger do? change the ownership matrix.

- Step 1: Recover marginal costs $\widehat{\mathbf{mc}} = \mathbf{p} + \Omega(\mathbf{p})^{-1} q(\mathbf{p})$.
- Step 1a: (Possibly) adjust marginal cost $\widehat{\mathbf{mc}} \cdot (1 - e)$ with some cost efficiency e.
- Step 2: Change the ownership matrix $\Omega^{pre}(\mathbf{p}) \rightarrow \Omega^{post}(\mathbf{p})$.
- Step 3: Solve for $\mathbf{p}^{post}$ via: $\mathbf{p} = \widehat{\mathbf{mc}} - \Omega(\mathbf{p})^{-1} q(\mathbf{p})$.

Merger Simulation

What does a merger do? change the ownership matrix.

- Step 1: Recover marginal costs $\widehat{\mathbf{mc}} = \mathbf{p} + \Omega(\mathbf{p})^{-1} q(\mathbf{p})$.
- Step 1a: (Possibly) adjust marginal cost $\widehat{\mathbf{mc}} \cdot (1 - e)$ with some cost efficiency e.
- Step 2: Change the ownership matrix $\Omega^{pre}(\mathbf{p}) \rightarrow \Omega^{post}(\mathbf{p})$.
- Step 3: Solve for $\mathbf{p}^{post}$ via: $\mathbf{p} = \widehat{\mathbf{mc}} - \Omega(\mathbf{p})^{-1} q(\mathbf{p})$.

- The first step is easy (just a matrix inverse).
- The second step is trivial.
- The third step is tricky because we have to solve an implicit system of equations. $\mathbf{p}$ is on both sides.

Solution Methods

We were a bit vague on how we were solving the system of equations for p^{post}, the post-merger price.

General problem $F(x) = 0$ or n nonlinear equations and n unknowns $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

$$\begin{aligned} F_1(x_1, \ldots, x_n) &= 0 \\ F_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ F_{N-1}(x_1, \ldots, x_n) &= 0 \\ F_N(x_1, \ldots, x_n) &= 0 \end{aligned}$$

Solution Methods

Helpful to write $F(x) = 0 \Leftrightarrow x - \alpha F(x) = x$ which yields the fixed point problem:

$$G(x) = x - \alpha F(x)$$

Fixed point iteration

$$x^{k+1} = G(x^k)$$

Nonlinear Richardson iteration or Picard iteration.

We need G to be a contraction mapping for iterative methods to guarantee a unique solution (often need strong monotonicity as well).

Gauss Jacobi: Simultaneous Best Reply

Current iterate: $x^k = (x_1^k, x_2^k, \ldots, x_{n-1}^k, x_n^k)$.

Compute the next iterate x^{k+1} by solving one equation in one variable using only values from x^k :

$$
\begin{aligned}
F_1(x_1^{k+1}, x_2^k \ldots, x_{n-1}^k, x_n^k) &= 0 \\
F_2(x_1^k, x_2^{k+1}, \ldots, x_{n-1}^k, x_n^k) &= 0 \\
&\vdots \\
F_{n-1}(x_1^k, x_2^k, \ldots, x_{n-1}^k, x_n^k) &= 0 \\
F_n(x_1^k, x_2^k, \ldots, x_{n-1}^k, x_n^k) &= 0
\end{aligned}
$$

Requires contraction and strong monotonicity.

Partial Merger Analysis

- Hold all other prices p_{-j} fixed at pre-merger prices.
- Adjust the marginal costs for potential efficiencies.
- Consider only the FOC for product j

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

- Solve for the new p_j given the change in the products controlled by firm f: $\mathcal{J}_f \to \mathcal{J}_f'$
- This is a single Gauss-Jacobi step (only products involved in merger).

Partial Merger Analysis: Why bother?

- We only need own and cross elasticities for products involved in the merger.
- Tends to show smaller price increases than full equilibrium merger analysis.
- Only solving a single equation rather than a system of J nonlinear equations.

Gauss Seidel: Iterated Best Response

Current iterate: $x^k = (x_1^k, x_2^k, \ldots, x_{n-1}^k, x_n^k)$.

Compute the next iterate x^{k+1} by solving one equation in one variable updating as we go through:

$$
\begin{array}{rcl}
F_1(x_1^{k+1}, x_2^k \ldots, x_{n-1}^k, x_n^k) & = & 0 \\
F_2(x_1^{k+1}, x_2^{k+1}, \ldots, x_{n-1}^k, x_n^k) & = & 0 \\
& \vdots & \\
F_{n-1}(x_1^{k+1}, x_2^{k+1}, \ldots, x_{n-1}^{k+1}, x_n^k) & = & 0 \\
F_n(x_1^{k+1}, x_2^{k+1}, \ldots, x_{n-1}^{k+1}, x_n^{k+1}) & = & 0
\end{array}
$$

Requires contraction and strong monotonicity.

Newton's Method

1. Take an initial guess x^0
2. Take a Newton step by solving the following system of linear equations

$$J_F(x^k)s^k = -F(x^k)$$

3. New guess $x^{k+1} = x^k + s^k$.
4. Good (Quadratic) Local convergence

- Requires J_F (Jacobian) to be Lipschitz continuous.
- Linearity means we do not need to take the inverse to solve the system (just QR decomp – `backslash` in MATLAB).
- Non-singularity of J_F is weaker than strong monotonicity (more like PSD).

Fixed Point Iteration?

- Can we iterate on the price relation until we converge to a new equilibrium?

$$\mathbf{p} \leftarrow \widehat{\mathbf{mc}} - \Omega(\mathbf{p})^{-1} q(\mathbf{p})$$

- While tempting, this doesn't work. (It is not a contraction).
- There is a modification that is a contraction for logit type models.
- You can always get lucky(!)

Morrow Skerlos (2010) Fixed Point

- For the logit (and variants) we can factor $\frac{\partial q_j}{\partial p_k}$ into two parts.

$$\Omega_{jk}(\mathbf{p}) = \underbrace{\alpha \cdot I[j=k] \cdot s_j(\mathbf{p})}_{\Lambda(\mathbf{p})} - \underbrace{\alpha \cdot s_j(\mathbf{p}) s_k(\mathbf{p})}_{\Gamma(\mathbf{p})}$$

- $\Gamma(\mathbf{p})$ and $\Lambda(\mathbf{p})$ are $J \times J$ matrices and $\Lambda(\mathbf{p})$ is diagonal and (j,k) is nonzero in $\Gamma(\mathbf{p})$ only if (j,k) share an owner.
- After factoring we can rescale by $\Lambda^{-1}(\mathbf{p})$

$$(\mathbf{p} - \mathbf{mc}) \leftarrow \Lambda^{-1}(\mathbf{p}) \cdot \Gamma(\mathbf{p}) \cdot (\mathbf{p} - \mathbf{mc}) - \Lambda^{-1}(\mathbf{p}) \cdot s(\mathbf{p})$$

- This alternative fixed point is in fact a contraction.
- Moreover the rate of convergence is generally fast and stable (much more than Gauss-Seidel or Gauss-Jacobi).

What are Diversion Ratios?

Horizontal Merger Guidelines (2010 rev.)

In some cases, the Agencies may seek to quantify the extent of direct competition between a product sold by one merging firm and a second product sold by the other merging firm by estimating the diversion ratio from the first product to the second product. The diversion ratio is the fraction of unit sales lost by the first product due to an increase in its price that would be diverted to the second product. Diversion ratios between products sold by one merging firm and products sold by the other merging firm can be very informative for assessing unilateral price effects, with higher diversion ratios indicating a greater likelihood of such effects. Diversion ratios between products sold by merging firms and those sold by non-merging firms have at most secondary predictive value.

This time with equations

Raise price of good j. People leave. What fraction of leavers switch to k?

$$D_{jk}(p_j, p_{-j}) = \frac{\frac{\partial q_k}{\partial p_j}}{\left|\frac{\partial q_j}{\partial p_j}\right|}$$

It's one of the best ways economists have to characterize competition among sellers.

- High Diversion: Close Substitutes $\rightarrow$ Mergers more likely to increase prices.
- Very low diversion $\rightarrow$ products may not be in the same market.
  (ie: Katz & Shapiro). This is just hypothetical monopolist or SSNIP test.
- Demand Derivatives NOT elasticities.
- No equilibrium responses.

Unilateral Effects

- Eliminating competition between the merging firms can itself constitute a substantial lessening of competition
- Developed in the 1992 Guidelines, and larger role in the 2010 Guidelines
- Based on modern theoretical literature: Farrell Shaprio (1990), Werden (1996), Farrel Shapiro (2010), Froeb and Werden (1998)
- Extension to multiple products/firms may be tricky (Carlton 2010, Hausman, Moresei, Rainey (2010)).
- Doesn't go as far as pass-through literature (Bulow Geanakoplos Klemperer (1985), Jaffe Weyl (2013)).
- Limited empirical results in academic literature: (Cheung 2013, Miller, Remer, Ryan, Sheu (2013), Conlon Mortimer (2013/5))
- Possibly more empirical experience at DOJ/FTC.

Where do Diversion Ratios come from?
(Stolen from Conlon and Mortimer (2018)

In Theory

Consider Bertrand FOC's for single-product firm j, buys k:

$$
\begin{aligned}
\arg\max_{p_j} \quad & (p_j - c_j) \cdot q_j(p_j, p_{-j}) + (p_k - c_k) \cdot q_k(p_j, p_{-j}) \\
0 &= q_j + (p_j - c_j) \cdot \frac{\partial q_j}{\partial p_j} + (p_k - c_k) \cdot \frac{\partial q_k}{\partial p_j} \\
p_j &= -q_j / \frac{\partial q_j}{\partial p_j} + c_j + (p_k - c_k) \cdot \underbrace{\frac{\partial q_k}{\partial p_j} / - \frac{\partial q_j}{\partial p_j}}_{D_{jk}}
\end{aligned}
$$

In Theory

Consider Bertrand FOC's for single-product firm j, buys k:

$$
\begin{aligned}
\arg\max_{p_j} \quad & (p_j - c_j) \cdot q_j(p_j, p_{-j}) + (p_k - c_k) \cdot q_k(p_j, p_{-j}) \\
0 \quad = \quad & q_j + (p_j - c_j) \cdot \frac{\partial q_j}{\partial p_j} + (p_k - c_k) \cdot \frac{\partial q_k}{\partial p_j} \\
p_j \quad = \quad & -q_j / \frac{\partial q_j}{\partial p_j} + c_j + (p_k - c_k) \cdot \underbrace{\frac{\partial q_k}{\partial p_j} / - \frac{\partial q_j}{\partial p_j}}_{D_{jk}} \\
p_j \quad = \quad & \underbrace{\frac{\epsilon_{jj}}{\epsilon_{jj} + 1}}_{\text{Lerner Markup}} \left[c_j - \overbrace{\underbrace{c_j \cdot e_j}_{\text{efficiency}} + \underbrace{(p_k - c_k) \cdot D_{jk}(p_j, p_{-j})}_{\text{opp cost}}}^{UPP} \right]
\end{aligned}
$$

Caveat: UPP, Partial Merger, Full Merger.

UPP Extensions

Extension to multiple acquisitions:
Very easy if we have that $p_j - mc_j = p - mc$ are the same for several values of j. Then

$$UPP_j \approx (p - mc) \sum_k D_{jk}(\mathbf{p}) - E_j mc_j$$

If several brands of acquisition have the same markup – can consider firm-level diversion. (We can aggregate diversion across similar flavors)

Ignoring Efficiencies

$$GUPPI_j \approx \frac{(p_j - mc_j)}{p_j} D_{jk}(\mathbf{p})$$

Diversion: In Practice

1. Calculated from an estimated demand system (ratio of estimated cross-price to own-price demand derivatives)
2. Consumer surveys (what would you buy if not this?)
3. Obtained in 'course of business' (sales reps, internal reviews)

Antitrust authorities may prefer different measures in different settings. Are they concerned about:

- Small but widespread price hikes?
- Product discontinuations or changes to availability?

Is it sufficient to rely on data from merging firms only?

- Do we need diversion to other products in the 'market' or other functions of market-level data?
- Discrete-choice demand models imply that 'aggregate diversion' (including to an outside good) sums to one.

Diversion has treatment effects interpretation

Treatment	"not purchasing j"
Outcome	fraction of j consumers who switch to product k
Treated group	consumers who would have purchased j at pre-merger price, but do not purchase at a higher price

Heterogeneity: Individuals who leave j after a \$0.01 price increase differ in their taste for k from those who leave after \$1, \$100, \$10,000 price increases.

Start with the Wald Estimator

Consider an experiment designed to measure diversion, where everything else is held fixed and p_j is exogenously increased by Δp_j:

$$
\begin{aligned}
D_{jk}(p_j, p_{-j}) &= \left| \frac{q_k(p_j + \Delta p_j, p_{-j}) - q_k(p_j, p_{-j})}{q_j(p_j + \Delta p_j, p_{-j}) - q_j(p_j, p_{-j})} \right| \\
&= \frac{\int_{p_j^0}^{p_j^0 + \Delta p_j} \frac{\partial q_k(p_j, p_{-j})}{\partial p_j} \partial p_j}{\Delta q_j}
\end{aligned}
$$

Re-write as Local Average Treatment Effect

$$\widehat{D_{jk}}^{LATE} = \frac{1}{\Delta q_j} \int_{p_j^0}^{p_j^0+\Delta p_j} D_{jk}(p_j, p_{-j}) \left| \frac{\partial q_j(p_j, p_{-j})}{\partial p_j} \right| \partial p_j$$

- $\widehat{D_{jk}}^{LATE}$ is a Local Average Treatment Effect (LATE).
  - Identified from finite price changes (simulated or actual).
  - For any finite price increase, we measure a weighted average of the diversion function, where the weights are the lost sales of j: $w(\mathbf{p}) = \frac{1}{\Delta q_j} \frac{\partial q_j(p_j, p_{-j})}{\partial p_j}$

Re-write as Local Average Treatment Effect

$$\widehat{D_{jk}}^{LATE} = \frac{1}{\Delta q_j} \int_{p_j^0}^{p_j^0 + \Delta p_j} D_{jk}(p_j, p_{-j}) \left| \frac{\partial q_j(p_j, p_{-j})}{\partial p_j} \right| \partial p_j$$

- $\widehat{D_{jk}}^{LATE}$ is a Local Average Treatment Effect (LATE).
    - Identified from finite price changes (simulated or actual).
    - For any finite price increase, we measure a weighted average of the diversion function, where the weights are the lost sales of j: $w(\mathbf{p}) = \frac{1}{\Delta q_j} \frac{\partial q_j(p_j, p_{-j})}{\partial p_j}$
- Let $\widehat{D_{jk}}^{ATE}$ denote Average Treatment Effect (ATE) when everyone is treated.
    - Δp_j increases to choke price: $Q_j(p_j^0 + \Delta p_j, p_{-j}) = 0$.
    - Interpretation as second-choice data.

The Nonparametric Object: MTE

Re-writing:

$$\widehat{D_{jk}}^{LATE}(p_j, p_{-j}) = \frac{1}{\Delta q_j} \int_{p_j^0}^{p_j^0 + \Delta p_j} D_{jk}(p_j, p_{-j}) \left| \frac{\partial q_j(p_j, p_{-j})}{\partial p_j} \right| \partial p_j$$

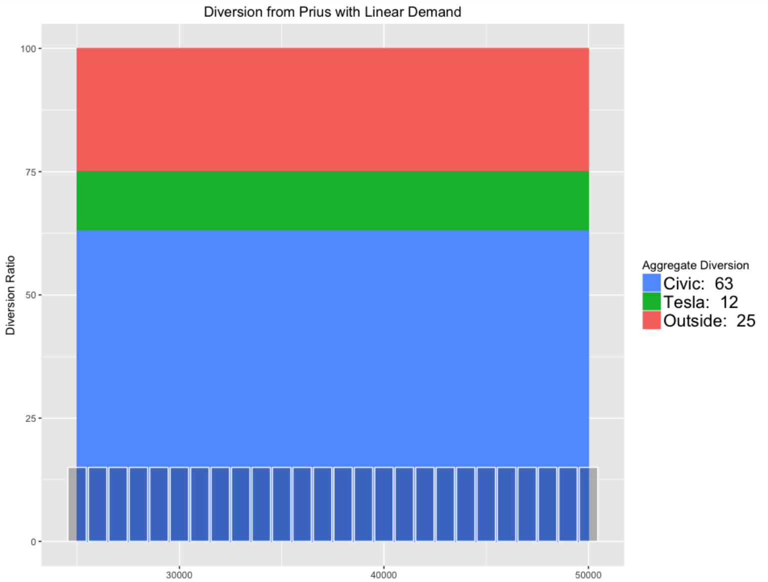
- Diversion, $D_{jk}(p_j, p_{-j})$, is a Marginal Treatment Effect (MTE) in the language of Heckman and Vytlacil (1999).
- It is a function. Actually a matrix valued function.
- It is not identified non-parametrically from a single price increase.

Various Treatment Effects

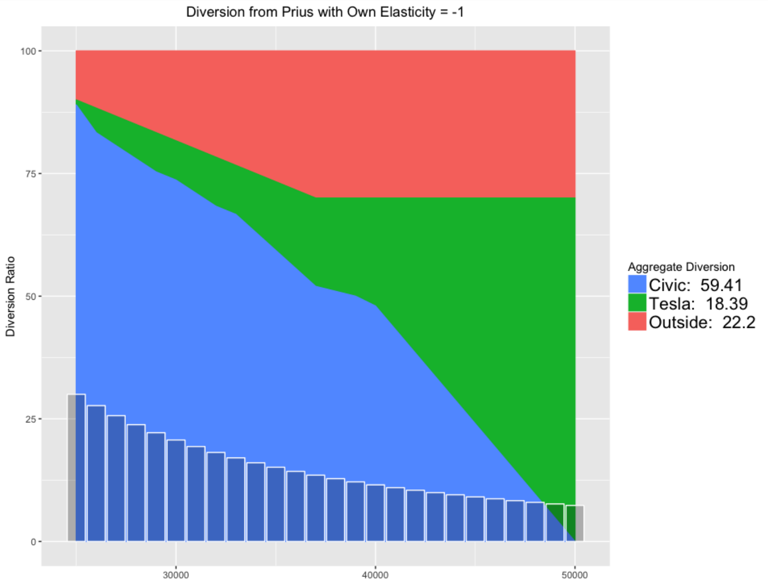
- Determine what different measures of diversion identify.
    - Finite price increase $\rightarrow$ local average treatment effect (LATE)
    - Product removal (treating everyone) $\rightarrow$ average treatment effect (ATE)
    - A nonparametric function of p_j $\rightarrow$ marginal treatment effect (MTE)
    - Constant diversion: three measures coincide (Theory/Empirics)

But... How do the weights work? An illustration.

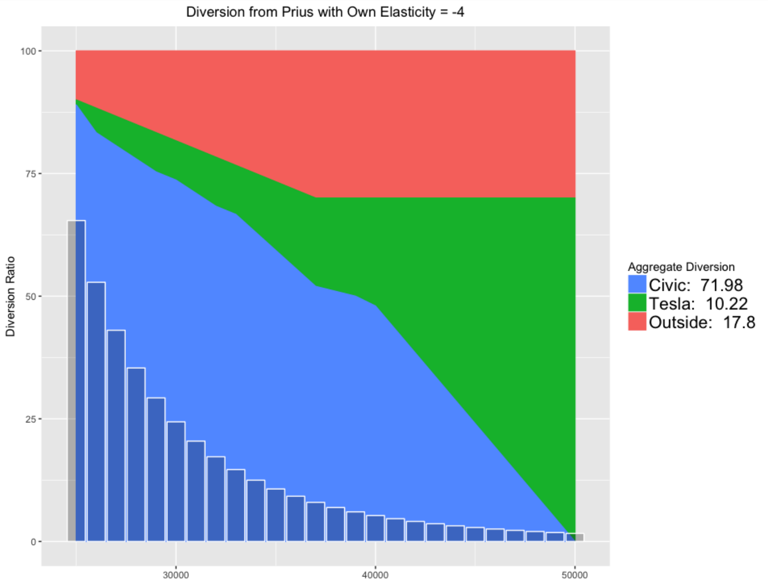
Thought Experiment – Linear Demand for a Toyota Prius



Thought Experiment – Inelastic CES Demand for a Prius



Thought Experiment – Elastic CES Demand for a Prius



Bias of Estimator

How far apart are $D_{jk}(\mathbf{p^0})$ and $\widehat{D_{jk}}$ when we increase price by Δp_j?

$$\begin{aligned}
q_k(\mathbf{p} + \Delta p_j) &\approx q_k(\mathbf{p}) + \frac{\partial q_k}{\partial p_j}\Delta p_j + \frac{\partial^2 q_k}{\partial p_j^2}(\Delta p_j)^2 + O((\Delta p_j)^3) \\
\frac{q_k(\mathbf{p} + \Delta p_j) - q_k(\mathbf{p})}{\Delta p_j} &\approx \frac{\partial q_k}{\partial p_j} + \frac{\partial^2 q_k}{\partial p_j^2}\Delta p_j + O(\Delta p_j)^2 \\
Bias(\widehat{D_{jk}} - D_{jk}(\mathbf{p^0})) &\approx -\frac{D_{jk}\frac{\partial^2 q_j}{\partial p_j^2} + \frac{\partial^2 q_k}{\partial p_j^2}}{\frac{\partial q_j}{\partial p_j} + \frac{\partial^2 q_j}{\partial p_j^2}\Delta p_j}\Delta p_j
\end{aligned}$$

- The downside of a large change Δp_j is that the approximation of demand at $\mathbf{p^0}$ is less accurate and depends on the curvature of the demand function.
- There are two demand models for which $Bias \equiv 0$ (constant treatment effects):

Bias of Estimator

How far apart are $D_{jk}(\mathbf{p^0})$ and $\widehat{D_{jk}}$ when we increase price by Δp_j?

$$\begin{aligned}
q_k(\mathbf{p}+\Delta p_j) &\approx q_k(\mathbf{p}) + \frac{\partial q_k}{\partial p_j}\Delta p_j + \frac{\partial^2 q_k}{\partial p_j^2}(\Delta p_j)^2 + O((\Delta p_j)^3) \\
\frac{q_k(\mathbf{p}+\Delta p_j) - q_k(\mathbf{p})}{\Delta p_j} &\approx \frac{\partial q_k}{\partial p_j} + \frac{\partial^2 q_k}{\partial p_j^2}\Delta p_j + O(\Delta p_j)^2 \\
Bias(\widehat{D_{jk}} - D_{jk}(\mathbf{p^0})) &\approx -\frac{D_{jk}\frac{\partial^2 q_j}{\partial p_j^2} + \frac{\partial^2 q_k}{\partial p_j^2}}{\frac{\partial q_j}{\partial p_j} + \frac{\partial^2 q_j}{\partial p_j^2}\Delta p_j}\Delta p_j
\end{aligned}$$

- The downside of a large change Δp_j is that the approximation of demand at $\mathbf{p^0}$ is less accurate and depends on the curvature of the demand function.
- There are two demand models for which $Bias \equiv 0$ (constant treatment effects): linear demand and plain IIA logit.

Variance of Estimator

$$Var(\widehat{D_{jk}}) \quad \approx \quad Var\left(\frac{\Delta q_k}{|\Delta q_j|}\right) \approx \frac{D_{jk}(1-D_{jk})}{\Delta q_j} \approx \frac{D_{jk}(1-D_{jk})}{\left|\frac{\partial q_j}{\partial p_j}\right| \Delta p_j}$$

Variance is a problem when:

- $\left|\frac{\partial q_j}{\partial p_j}\right|$ is small (inelastic demand $\rightarrow$ market power/ when we may be most worried about mergers).
- $\Delta p_j \approx 0$ (when price change is small).
- Exacerbated by variation in (q_j, q_k) unrelated to the exogenous price change (stochastic demand).

Bias-Variance tradeoff

- Precise measure of $\widehat{D_{jk}}^{ATE}$ or $\widehat{D_{jk}}^{LATE}$ for a large Δp_j vs.
- Noisy measure of $D_{jk}(\mathbf{p^0})$

Nevo (2000) and BLP (1999) Applications

Data from Nevo (2000): $T = 94$ markets, $J = 24$ brands.

- RTE cereal (e.g., Kellogg's and General Mills merger)

$$u_{ijt} = d_j + x_{jt} \underbrace{\left(\overline{\beta} + \Sigma \cdot \nu_i + \Pi \cdot d_{it}\right)}_{\beta_{it}} + \Delta\xi_{jt} + \varepsilon_{ijt}$$

  - Features a large amount of preference heterogeneity, especially with respect to the price sensitivity β_{it}^{price}
  - Estimated coefficient on price is distributed:

$$\beta_{it}^{price} \sim N\left(\text{-63 + 588} \cdot \text{income}_{it} \text{ - 30} \cdot \text{inc}_{it}^{2} + 11 \cdot \text{I[child]}_{it}, \sigma\text{=3.3}\right)$$

Data from BLP (1999): $T = 21$ markets, $J \approx 150$ products per market (total of 2271 product-market pairs)

- Random coefficients on vehicle size, miles-per-dollar, AC, horsepower/weight, constant. Price coefficient depends on income.

Nevo (2000) and BLP (1999) Applications, cont.

Define:

$$MTE = \frac{\frac{\partial s_k}{\partial p_j}}{\left|\frac{\partial s_j}{\partial p_j}\right|}, \quad ATE = \frac{s_k(A \setminus j) - s_k(A)}{|s_j(A \setminus j) - s_j(A)|}, \quad Logit = \frac{s_k(A)}{1 - s_j(A)}$$

- Compare $MTE(\mathbf{p_0})$ to ATE
- Compare $MTE(\mathbf{p_0})$ to Logit (Constant diversion, $\propto$ to share.)

Nevo (2000) Results

Three Measures of Diversion

	MTE	*ATE*	*Logit*
	Best Substitute		
Med(D_{jk})	13.26	13.54	9.05
Mean(D_{jk})	15.11	15.62	10.04
% Agree with MTE		89.98	58.38
	Outside Good		
Med(D_{j0})	35.30	32.40	54.43
Mean(D_{j0})	36.90	33.78	53.46

The first panel reports diversion to each product-market pair's best substitute. The second panel reports diversion to the outside good.

BLP (1999) Results

	MTE	ATE	$Logit$
	Best Substitute		
Med(D_{jk})	5.10	5.04	0.46
Mean(D_{jk})	6.07	6.25	0.53
% Agree with MTE	100.00	96.89	95.62
	Outside Good		
Med(D_{j0})	17.05	13.02	89.26
Mean(D_{j0})	17.04	13.44	89.36

The first panel reports diversion to each product-market pair's best substitute. The second panel reports diversion to the outside good.

Nevo (2000) Results, cont.

% Difference in Diversion Measures: y vs. $x = \log(D\widehat{^{MTE}(\mathbf{p_0}})$)

	med$(y - x)$	mean$(y - x)$	med$(\|y - x\|)$	mean$(\|y - x\|)$	std$(\|y - x\|)$
			Best Substitutes		
ATE	2.56	3.24	6.00	7.61	7.04
$Logit$	-44.19	-42.88	44.92	47.77	28.63
			All Products		
ATE	5.78	8.30	8.29	12.13	12.02
$Logit$	-35.90	-25.92	49.48	53.27	34.56
			Outside Good		
ATE	-7.93	-8.89	7.94	9.08	6.77
$Logit$	39.22	39.20	39.22	40.60	22.05

Table compares ATE and Logit measures of diversion to the MTE measure.

The first panel reports differences for each product-market pair's best substitute.

The second panel averages across all possible substitutes.

BLP (1999) Results, cont.

% Difference in Diversion Measures: y vs. $x = \log(D^{\widehat{MTE}}(\mathbf{p_0}))$

	med$(y-x)$	mean$(y-x)$	med$(\|y-x\|)$	mean$(\|y-x\|)$	std$(\|y-x\|)$
			Best Substitutes		
ATE	-0.53	0.08	11.51	12.64	9.76
$Logit$	-232.16	-239.75	232.16	239.75	40.58
			All Products		
ATE	9.79	26.52	22.54	40.34	47.85
$Logit$	-183.79	-162.21	186.39	177.35	86.11
			Outside Good		
ATE	-23.62	-24.25	23.67	24.99	13.40
$Logit$	165.42	186.43	165.42	186.43	72.86

Lessons from Nevo (2000) and BLP (1999)

- MTE vs. ATE measures are not hugely different in Nevo (2000).
- Larger differences in BLP (1999). Why? More variation in quality, cost, and especially price – better opportunity to observe larger differences in diversion.
- ATE tends to predict slightly more inside substitution and less outside substitution.
- ATE may either overstate or understate diversion to other products on average. If the marginal consumer is more (less) inelastic as price increases, then ATE over- (under-) states diversion.
- Both models rely on sum of diversion $= 1$.
- Imposing proportional substitution (Logit) looks terrible.

What's the point/Extensions

- Calculating D_{jk} or the matrix gives us the best idea about which products compete with each other.
- What is wrong with cross-price elasticities?
- Can we go from opportunity costs to prices?

Conduct

Conduct Overview

- A second set of important questions in IO is being able to use data to decide whether firms are competing or colluding.
- Absent additional restrictions, we cannot generally look at data on (P, Q) and decide whether or not collusion is taking place.
- 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)

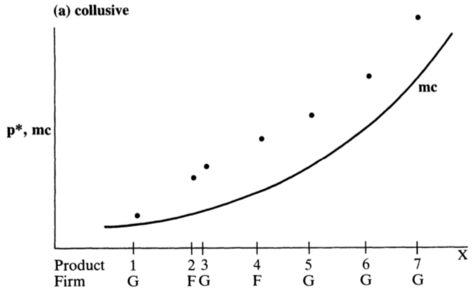


Figure 2(a)

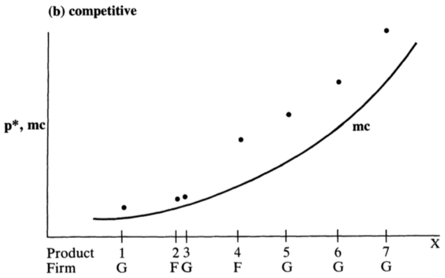


Figure 2(b)

Testing For Conduct: Challenges

- Recall the Ω matrix which we can write as $\Omega = \tilde{\Omega} \odot A$, where $\odot$ is the element-wise or Hadamard product of two matrices.
    - $\tilde{\Omega}$ is the matrix of demand derivatives with $\Omega_{(j,k)} = \frac{\partial q_j}{\partial p_k}$ for all elements.
    - $A_{(j,k)} = 1$ if (j,k) have the same owner and 0 otherwise.
- Mergers are about changing 0's to 1's in the A matrix.
- Matrix form of FOC: $q(\mathbf{p}) = \Omega(\mathbf{p}) \cdot (\mathbf{p} - \mathbf{mc})$

Testing For Collusion: Challenges

We derived those conditions from multi-product Bertrand FOCs:

$$\begin{aligned}
\arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p}) &= \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot q_j(\mathbf{p}) + \kappa_{fg} \sum_{j \in \mathcal{J}_g} (p_j - c_j) \cdot q_j(\mathbf{p}) \\
\rightarrow 0 &= q_j(\mathbf{p}) + \sum_{k \in (\mathcal{J}_f, \mathcal{J}_g)} \kappa_{fg} \cdot (p_k - c_k) \frac{\partial q_k}{\partial p_j}(\mathbf{p})
\end{aligned}$$

- Now we have generalized the $A(\kappa)$ matrix.
- 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 colluded) 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.

Reasons for Deviations from Static Bertrand

Biased estimates of own and cross price derivatives: For anything to work, you have correct estimates of $\tilde{\Omega}$. My prior is most papers underestimate cross price elasticites.

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.

Algorithm #1: Bertrand Deviations

- Recover $\tilde{\Omega}$ from demand alone.
- Recover marginal costs $\widehat{\mathbf{mc}} = \mathbf{p} + (O. * \tilde{\Omega}(\mathbf{p}))^{-1} q(\mathbf{p})$.

Challenges:

- Given $[\mathbf{q}, \mathbf{p}, \tilde{\Omega}, O]$ I can always produce a vector of marginal costs $\mathbf{c}$ that rationalizes what we observe. [ie: J equations J unknowns].
- Maybe some vectors of $\mathbf{c}$ look less "reasonable" than others.
  - ie: I have a parametric model of MC in mind.
  - Can test that model with GMM objective of c_{jt} on regressors.
  - Maybe marginal costs cannot deviate too much within product from period to period.
  - Marginal costs ≤ 0 seem problematic. [Might just be that your estimates for demand are too inelastic...]

Algorithm #2: Simultaneous Supply and Demand

- Recover $\tilde{\Omega}$ from demand and parametric assumption on supply (GMM with both sets of moments).
- I can impose $c > 0$ by using $\ln mc_{jt} = x_{jt}\gamma_1 + w_{jt}\gamma_2 + \omega_{jt}$.
- 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].

Challenges:

- Am I testing conduct? Or am I testing the linear functional form for my supply model?
- Will a missing z_{jt} change whether or not I believe firms are colluding?

Algorithm #3: Exclusion Restrictions

- We provide a formal test for four alternative models of conduct based on the exclusion restriction test in Berry and Haile (2014)

$$
\begin{aligned}
\widehat{mc_{jt}}(\kappa, \hat{\theta}) &= \lambda_j + \gamma_1 x_{jt} + \gamma_2 w_{jt} + \omega_{jt} \\
\omega_{jt} &= \widehat{mc_{jt}}(\kappa, \hat{\theta}) - \lambda_j - \gamma_1 x_{jt} - \gamma_2 w_{jt} \\
0 &= E[\omega_{jt} | \lambda_j, x_{jt}, w_{jt}, z^s_{jt}]
\end{aligned}
$$

- w_{jt}: cost shifters (price of corn for Corn Flakes, price of rice for Rice Krispies).
- z^s_{jt}: should not shift marginal costs under the true model of conduct but could potentially shift marginal costs under the alternative. A good choice is markup shifters.
  - BLP instruments
  - Cost shifters for other products (Price of Rice for Corn Flakes, Price of Corn for Rice Krispies).
  - κ parameters or κ weighted diversion.

Start with BLP(95/99) / Nevo (2001)

Utility of consumer i for product j and store-week t as:

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}$$

Market shares are given by:

$$s_{jt}(\delta_{\cdot t}, \theta_2) = \int \frac{\exp[\delta_{jt} + \mu_{ijt}]}{\sum_{k \in J_t} \exp[\delta_{kt} + \mu_{ikt}]} f(\mu_{it} \mid \widetilde{\theta}_2)\, d\mu_{it}.$$

BH2014 show that one can invert the vector of observed market shares $\mathcal{S}_t$ to solve for $\delta_t = D_t^{-1}(\mathcal{S}_t, \theta_2)$.

Supply Side

Consider the multi-product Bertrand FOCs:

$$
\begin{aligned}
\arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p}) &= \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot s_j(\mathbf{p}) + \kappa_{fg} \sum_{k \in \mathcal{J}_g} (p_k - c_k) \cdot s_k(\mathbf{p}) \\
0 &= s_j(\mathbf{p}) + \sum_{k \in \mathcal{J}_f} (p_k - c_k) \frac{\partial s_k}{\partial p_j}(\mathbf{p}) + \sum_g \kappa_{fg} \sum_{l \in \mathcal{J}_g} (p_l - c_l) \frac{\partial s_l}{\partial p_j}(\mathbf{p})
\end{aligned}
$$

It is helpful to define the matrix $\Omega_{(j,k)}(\mathbf{p}) = -\frac{\partial s_j}{\partial p_k}(\mathbf{p})$:

$$
A(\kappa)_{(j,k)} = \left\{ \begin{array}{ll} 1 & \text{for } j \in \mathcal{J}_f \\ \kappa_{fg} & \text{for } j \in \mathcal{J}_f, k \in \mathcal{J}_g \\ 0 & \text{o.w} \end{array} \right\}
$$

We can re-write the FOC in matrix form:

$$
\begin{aligned}
s(\mathbf{p}) &= (A(\kappa) \odot \Omega(\mathbf{p})) \cdot (\mathbf{p} - \mathbf{mc}), \\
\mathbf{mc} &= \mathbf{p} - \underbrace{(A(\kappa) \odot \Omega(\mathbf{p}))^{-1} s(\mathbf{p})}_{\eta(\mathbf{p},\mathbf{s},\theta_2,\kappa)}.
\end{aligned}
$$

Simultaneous Problem

Assume additivity, and write in terms of structural errors:

$$\xi_{jt} = \delta_{jt}(\mathcal{S}_t, \widetilde{\theta}_2) - \theta_1[x_{jt}, v_{jt}] - \alpha p_{jt}$$
$$\omega_{jt} = f\left(p_{jt} - \eta_{jt}(\mathbf{p}, \mathbf{s}, \theta_2, \kappa)\right) - h(x_{jt}, w_{jt}, \theta_3)$$

We've highlighted the two exclusion restrictions:

- Cost shifters w_{jt}
- Demand shifters v_{jt}

To simplify slides we let $f(x) = x$ (often $f(x) = \log(x)$).

Simultaneous Problem: Menu Approach

Assume two models of conduct (correct: κ_0) (incorrect: κ_1)

$$f(p_{jt} - \eta_{jt}(\kappa_0)) = h(x_{jt}, w_{jt}; \theta_3^0) + \omega_{jt}^0,$$
$$f(p_{jt} - \eta_{jt}(\kappa_1)) = h(x_{jt}, w_{jt}; \theta_3^1) + \omega_{jt}^1.$$

Write things in terms of the markup difference:

$$p_{jt} - \eta_{jt}(\kappa_1) = h(x_{jt}, w_{jt}; \theta_3) + \overbrace{\lambda \cdot \Delta\eta_{jt}(\mathbf{p}, \mathbf{s}, \theta, \kappa_0, \kappa_1) + \omega_{jt}}^{\widetilde{\omega_{jt}}}$$

Tempting idea: run the above regression and test if $\lambda = 0$.

- True model $\lambda = 0$, alternate model $\lambda \neq 0$.

Simultaneous Problem: Menu Approach

Assume two models of conduct (correct: κ_0) (incorrect: κ_1)

$$f(p_{jt} - \eta_{jt}(\kappa_0)) = h(x_{jt}, w_{jt}; \theta_3^0) + \omega_{jt}^0,$$
$$f(p_{jt} - \eta_{jt}(\kappa_1)) = h(x_{jt}, w_{jt}; \theta_3^1) + \omega_{jt}^1.$$

Write things in terms of the markup difference:

$$p_{jt} - \eta_{jt}(\kappa_1) = h(x_{jt}, w_{jt}; \theta_3) + \overbrace{\lambda \cdot \Delta\eta_{jt}(\mathbf{p}, \mathbf{s}, \theta, \kappa_0, \kappa_1) + \omega_{jt}}^{\widetilde{\omega_{jt}}}$$

Tempting idea: run the above regression and test if $\lambda = 0$.

- True model $\lambda = 0$, alternate model $\lambda \neq 0$.
- η_{jt} is endogenous: it depends on everything including (ξ, ω).

An Old Problem

- Bresnahan (1980/1982) recognized this problem: we need "rotations of demand".
- Most of the literature followed Bresnahan (1987):
  - ω_{jt} is measurement error in price
  - Ex: Bonnet and Dubois (2010) $E[\ln(\omega_{jt})|x_{jt}, w_{jt}] = 0$:

$$\log(p_{jt} - \eta_{jt}(\kappa, \widehat{\theta}_2))) = h(x_{jt}, w_{jt}, \theta_3) + \ln \omega_{jt}$$

- Other idea: put markup back on RHS and test $\lambda = 1$

$$p_{jt} = h(x_{jt}, w_{jt}, \theta_3) + \lambda \cdot \eta_{jt}(\kappa, \widehat{\theta}_2) + \omega_{jt}$$

  - "Informal" test of Villas Boas (2007): $E[\omega_{jt}|x_{jt}, w_{jt}] = 0$.
  - Pakes (2017) uses Wollman (2018) data and BLP IV $E[\omega_{jt}|x_{jt}, w_{jt}, f(x_{-j})] = 0$.

A subtle solution

- Berry Haile 2014 tell us we need marginal revenue shifters to act as exclusion restrictions.
- We need an instrument for $\Delta\eta_{jt}(\mathbf{p}, \mathbf{s}, \theta, \kappa_0, \kappa_1)$
  - Maybe not so hard since it is basically a function of everything.
  - Cannot have a direct effect on mc_{jt} (exclusion restriction).
- Idea would be to use $E[\omega_{jt}|x_{jt}, w_{jt}, z_{jt}^S] = 0$:

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

Candidate Instruments for z_{jt}^{s}

1. The demand shifter v_{jt}: maybe easy to find??
   - We use product recalls; prices of complements don't work so well.
2. BLP instruments $f(x_{-j})$: not always strong
   - Amit and JF have a nice paper showing how to choose $f(x_{-j})$
3. Can use the same logic to construct v_{-jt} or w_{-jt}
   - ie: cost shifters (or demand shifters) of competing goods.
   - Price of rice for Corn Flakes; price of corn for Rice Krispies.
   - Will depend on closeness of substitutes ΔPPI or D_{jk}.
4. Observed Conduct Shifters: κ_{fg}
   - Usually conduct is unobserved if we are testing it!
   - Index Inclusion Events (Fiona, Kennedy et. al); BlackRock-BGI Acquisition (AST)
   - Miller Weinberg (2017) use (pre/post merger for cartel participants).

Things that don't work

- ξ_{jt} only makes sense if you believe $Cov(\xi_{jt}, \omega_{jt}) = 0$.
- $p_{j,t,-s}$ (Hausman instruments) same good in other markets: pick up cost shocks (but could pick up changes in conduct!).
- If it isn't in one of our equations: does it have anything to do with demand or supply?
- It turns out that 2SLS analog $E[\Delta\eta_{jt}|x_t, w_t, v_t, Z^e_{jt}] = \widehat{\Delta\eta_{jt}}$ doesn't add much:
    - Markups aren't a linear function of observables.
    - Coefficients are (probably) quite different across products.