Common Ownership and Competition in the Ready-To-Eat Cereal Industry

Matthew Backus and Christopher Conlon and Michael Sinkinson
February 12, 2021
Introduction

1. Does the fact that firms have common investors lead to less competitive outcomes?

\[ \pi_f + \sum_{g \neq f} \kappa_{fg} \cdot \pi_g \quad \kappa_{fg} \geq 0 \]

- Managers maximize shareholder value, investors own portfolios including competitors → relaxes horizontal competition.
- Overlapping positions lead to intermediate case between own profit maximization and joint profit maximization.
- Growing (and controversial) literature:
  - We wrote a paper in *AEJ:Micro* (BCS Forthcoming) describing how ownership data maps into \( \kappa_{fg} \) and how the distribution looks for the broader economy.
  - Early lit has focused on price-concentration: Airlines Azar Schmalz Tecu (2018).
  - Anton, Ederer, Gine, Schmalz (2021) posit a plausible “quiet life” mechanism.
  - Alternatives: Boller and Scott Morton (2020): Index inclusion event study; Newham et. al Pharma Entry.
Introduction

2. Classic IO Q: How do we discern conduct (monopoly, PC, oligopoly, etc.) from observational data on \((P, Q)\)?

- Concerns:
  - If we could observe \(MC\) this would be pretty easy. (but mostly we cannot).
  - Many tests amount to joint test of conduct assumption (oligopoly, monopoly, perfect competition) and functional form of \(MC\) (linear, exponential, log-linear).
  - Different IV, weighting matrices, functional form assumptions may select different conduct assumptions.
• Take the nonparametric identification argument in Berry Haile (2014) and try to turn it into the most powerful (and general) semiparametric test.

• Use the testing framework of Rivers and Vuong (2002) (a LR type test) where:
  • Duarte, Magnolfi, Sullivan (2020) provide compelling evidence this is preferable to alternatives (e.g. Cox tests, Wald tests, etc.).
  • Null: Both models fit the data equally well. But both may be misspecified.

• A major focus is choosing good instruments that contain most of the information in the conditional moment restriction $E[\omega_{jt}|z_t] = 0$
  • What is the goal of IV? parallels to Chamberlain (1987). Choose $A(z_t)$ to maximize power (instead of efficiency).
  • Answer: Cheat and exploit model. Good IV predict the markup differences.
Setup and Assumptions

Assume we know demand $D(z_t)$ and define an additive markup $\eta_j(\cdot)$ as:

$$mc_{jt} \equiv p_{jt} - \eta_j(s_t, p_t, D(z_t))$$

$$mc_{jt} = h_s(x_{jt}, w_{jt}) + \omega_{jt}$$

$$\Rightarrow p_{jt} - \eta_{jt} = h_s(x_{jt}, w_{jt}) + \omega_{jt} \quad \text{where } E[\omega_{jt} | z_t] = 0$$

Assumptions

- Analogous to BH2014 $\eta_j(\cdot)$ is fully specified given demand.
  - e.g.: $\eta_t(p_t, s_t, D(z_t)) \equiv \Omega_t(p_t)^{-1}s_t(p_t)$
  - $\eta_{jt}^m$ where $m$ superscripts markup assumption (Cournot, Bertrand, Monopoly, PC)
- $\omega_{jt}$ is additively separable and existence of CMR $E[\omega_{jt} | z_t] = 0$
  - Prior work typically assumes $h_s(\cdot)$ linear or exponential ($\log mc_{jt}$)
- $\eta_{jt}$ is endogenous (depends on $\omega_t$) $\Rightarrow$ put on LHS (like AR test).

Goal: Choose the overidentifying restriction: $A(z_t)$. 

5
Testing Framework


- Estimation uses unconditional moments $\mathbb{E}[\omega_{jt}|z_t] = 0 \rightarrow \mathbb{E}[\omega_{jt}' A(z_t)] = 0$

\[
\frac{\sqrt{n}}{\sigma} \cdot (Q_W(\eta^1) - Q_W(\eta^2)) \overset{d}{\rightarrow} N(0, 1)
\]

- Idea: Both conditions can be violated $Q_W(\eta^m) > 0$.
- Prefer markup choice $\eta^m$ that leads to smaller violations (in GMM distance).
  - Calculating $\hat{\sigma}$ is often complicated $\rightarrow$ bootstrap.
  - Duarte, Magnolfi, Sullivan (2020) show RV outperforms Cox-type tests in simulation.
Proposition 1a: Standard GMM assumptions, fix $W$ and $h(x, w)$

$$Q_W(\eta^1) - Q_W(\eta^2) \xrightarrow{P} -\mathbb{E}[Z' \omega^1]' W \mathbb{E}[Z' \Delta \eta^{1,2}] - \mathbb{E}[Z' \omega^2]' W \mathbb{E}[Z' \Delta \eta^{1,2}]$$

- $\mathbb{E}[Z' \omega^1]$ and $\mathbb{E}[Z' \omega^2]$ are violation of moments (like we’d expect).
- $\mathbb{E}[Z' \Delta \eta^{1,2}]$ covariance of instruments with markup difference: “first stage”.

Proposition 1b: Under correct markup $\mathbb{E}[Z' \omega^1] \xrightarrow{a.s.} 0$

$$Q_W(\eta^1) - Q_W(\eta^2) \xrightarrow{P} -\mathbb{E}[Z' \Delta \eta^{1,2}] W \mathbb{E}[Z' \Delta \eta^{1,2}]$$

- Now just about correlation between instruments and $\Delta \eta^{1,2}$.
- When correlation is weak, models become indistinguishable.

Therefore we choose $A(z_t) = \mathbb{E}[\Delta \eta_{jt} | z_t]$. 
Procedure

Given demand and two markups $\eta_{jt}^1$ and $\eta_{jt}^2$ (e.g. perfect comp and monopoly):

1. Estimate $\hat{\omega}_{jt}^m$ as residual from (no IV necessary):

$$p_{jt} - \eta_{jt}^m = h_s(x_{jt}, w_{jt}) + \omega_{jt}^m$$

2. Estimate $\Delta \eta_{jt}^{1,2} = \hat{g}(z_t)$ as fitted value from (again no IV):

$$\Delta \eta_{jt}^{1,2} = g(z_t) + \zeta_{jt}$$

3. Compute the (scalar) moment violation: $\tilde{Q}(\eta^m) = \left(\frac{1}{N} \sum_{j,t} \hat{g}(z_t) \cdot \hat{\omega}_{jt}^m\right)^2$

4. Compare $T = \frac{\sqrt{n}}{\hat{\sigma}}(\tilde{Q}(\eta^1) - \tilde{Q}(\eta^2))$ to critical values of normal after estimating $\hat{\sigma}$ using bootstrap following Rivers and Vuong.

All regressions via random forest. (Note: different $\eta^1, \eta^2 \rightarrow$ different $A(z_t)$).
• Fully flexible $h_s(x_{jt}, w_{jt})$ (Don’t specify linear, log, etc.)
• Fully flexible $\Delta \eta_{jt}^{1,2} = g(z_t) + \zeta_{jt}$ ($z_t$ is very high dimensional).
• Random Forest is really good at complicated nonlinear forms.
• No weighting matrix $W$
• Theoretical analogue to optimal IV for “internalization parameter” [See paper].
• Easy to implement (fast enough to bootstrap).
Demand Estimation

- Discrete choice demand system based on BLP (1995), Nevo (2000/1) using Kilts Data:
  - Market: Chain-DMA-week (sampled 2/13 weeks per quarter)
  - Estimate market size from milk and egg purchases.
  - Correlated random coefficients on $p_{jt}$ and the constant.
  - 946 product FE and 1970 chain-dma-week FE.
- Demographics:
  - Chain-DMA-year specific demographics (income and children).
  - micro-moments matching income and children to price, characteristics (in PCA space), and outside good shares for 10 $\pi$ parameters.
- Instruments:
  - Own ingredient costs and chain specific demographic variables.
  - Quadratic Gandhi-Houde differentiation instruments
  - Calculate feasible approximation to optimal instruments (18): $E \left[ \frac{\partial \xi_{jt}}{\partial \theta} \mid Z_t \right]$.
- Estimation in PyBLP (Conlon and Gortmaker 2020).
Why RTE Cereal?

\[ C_4 \approx 85\% \text{ domestic, public firms and good ownership variation.} \]
Markups in dollars (Q4 2016)
# Counterfactual Mergers

<table>
<thead>
<tr>
<th>Firm</th>
<th>GM-KEL</th>
<th>Monopoly</th>
<th>$\kappa^{CO}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Mills</td>
<td>4.69</td>
<td>9.42</td>
<td>3.97</td>
</tr>
<tr>
<td>Kellogg’s</td>
<td>5.13</td>
<td>9.30</td>
<td>5.34</td>
</tr>
<tr>
<td>Quaker Oats</td>
<td>-0.37</td>
<td>14.87</td>
<td>7.75</td>
</tr>
<tr>
<td>Post</td>
<td>-0.15</td>
<td>12.76</td>
<td>7.06</td>
</tr>
<tr>
<td>Price Index</td>
<td>3.32</td>
<td>10.25</td>
<td>5.42</td>
</tr>
</tbody>
</table>

NB: Computed using marginal costs as predicted by own-profit maximization.
### Main Results: Assuming Linearity

<table>
<thead>
<tr>
<th>Own Profit Max vs.</th>
<th>Panel 1: $A(z_t) = z_t$, linear $h_s(\cdot)$; $W = (Z'Z)^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Ownership</td>
<td>-2.4732 -0.0079 -1.2333 -4.9099</td>
</tr>
<tr>
<td>Common Ownership (MA)</td>
<td>-2.5918 0.0070 -1.2105 -4.9215</td>
</tr>
<tr>
<td>Common Ownership (Lag)</td>
<td>-2.5208 0.0075 -1.2125 -4.9351</td>
</tr>
<tr>
<td>Perfect Competition</td>
<td>0.8611 -2.3033 -3.1652 -10.9229</td>
</tr>
<tr>
<td>Monopolist</td>
<td>-2.4166 -0.8783 -3.5162 -6.0048</td>
</tr>
</tbody>
</table>

| Own Profit Max vs. | Panel 2: $A(z_t) = \mathbb{E}[\Delta \eta_{jt}|z_t]$, linear $h_s(\cdot)$ and $g(\cdot)$ |
|-------------------|--------------------------------------------------------------------------------|
| Common Ownership  | -1.2859 -0.2126 -0.8317 -5.2361                                                |
| Common Ownership (MA) | -1.3993 -0.2071 -0.8340 -5.3019                                                |
| Common Ownership (Lag) | -1.3506 -0.2093 -0.8367 -5.3271                                               |
| Perfect Competition| 1.1732 -0.8843 -1.4708 -10.7559                                                |
| Monopolist        | -1.4038 -0.3243 -1.0613 -5.3183                                                |

Z-scores are reported. Bootstrap clustered by: Retailer-DMA-year

Predicting $\mathbb{E}[\Delta \eta_{jt}|z_t]$ is equivalent to a different choice of $W$. 

---

14
Main Results: Our (Semiparametric) Test

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Ownership</td>
<td>-4.8893</td>
<td>-5.4460</td>
<td>-5.4412</td>
<td>-5.9585</td>
</tr>
<tr>
<td>Common Ownership (MA)</td>
<td>-5.4345</td>
<td>-6.1348</td>
<td>-5.8757</td>
<td>-6.4357</td>
</tr>
<tr>
<td>Common Ownership (Lag)</td>
<td>-5.1770</td>
<td>-5.9221</td>
<td>-5.7041</td>
<td>-6.2255</td>
</tr>
<tr>
<td>Perfect Competition</td>
<td>-7.7749</td>
<td>-8.7051</td>
<td>-8.9758</td>
<td>-10.0654</td>
</tr>
<tr>
<td>Monopolist</td>
<td>-5.2711</td>
<td>-6.7789</td>
<td>-5.9158</td>
<td>-6.5933</td>
</tr>
</tbody>
</table>

Panel 3: \( A(z_t) = \mathbb{E}[\Delta_1^2 \eta_{12} | z_t] \), random forest \( h_s(\cdot) \) and \( g(\cdot) \)

Z-scores are reported. Bootstrap clustered by: Retailer-DMA-year

- Own-profit maximization wins by a landslide
- Choice of instruments doesn’t matter
- We capture the nonlinearity in \( h_s(\cdot), g(\cdot) \).
  - \( h_s(x_{jt}, w_{jt}) \) contains dummies for products and time periods, and own ingredient prices (e.g. corn for Corn Flakes), and product characteristics.
Internalization Parameters (Wald Approach)

Let $\kappa$ represent the weight a firm places on competitors and $\tau$ the internalization of those weights.

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

Now,

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

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

Our testing procedure has advantages over previous approaches:

- Amounts to two prediction exercises.
- We use the model itself to form $A(z_t)$.
- Flexible functional forms for $h_s(\cdot), g(\cdot)$ actually matter.
- No issues with weighting matrices.
- Nothing specific to common ownership.
- Anything that delivers a value for $\eta_{jt}$ is testable subject to relevance $E[z_t' \Delta \eta_{jt}]$.

No evidence of common ownership effects on prices in RTE Cereal.