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 $\rightarrow$ 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 κ_{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)?
   - History of identification: Bresnahan (1982), Lau (1982), Berry and Haile (2014)
   - Long history on testing conduct assumptions: Bresnahan (1987), Genesove and Mullin (1998), Nevo (1998/2001), Villas Boas (2007), Bonnet and Dubois (2010).
   - 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.

This Paper

- 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 $\mathbb{E}[\omega_{jt}|\boldsymbol{z_t}] = 0$
  - What is the goal of IV? parallels to Chamberlain (1987). Choose $A(\boldsymbol{z_t})$ to maximize power (instead of efficiency).
  - Answer: Cheat and exploit model. Good IV predict the markup differences.

Setup and Assumptions

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

$$
\begin{aligned}
mc_{jt} &\equiv p_{jt} - \eta_j(\boldsymbol{s_t}, \boldsymbol{p_t}, D(\boldsymbol{z_t})) \\
mc_{jt} &= h_s(x_{jt}, \mathrm{w}_{jt}) + \omega_{jt} \\
\rightarrow p_{jt} - \eta_{jt} &= h_s(x_{jt}, \mathrm{w}_{jt}) + \omega_{jt} \quad \text{where } \mathbb{E}[\omega_{jt} | \boldsymbol{z_t}] = 0
\end{aligned}
$$

Assumptions

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

Goal: Choose the overidentifying restriction: $A(\boldsymbol{z_t})$.

Testing Framework

Goal: non-nested model selection using Rivers Vuong (2002):

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

$$\frac{\sqrt{n}}{\sigma} \cdot (Q_W(\eta^1) - Q_W(\eta^2)) \xrightarrow{d} N(0, 1)$$

- Idea: Both conditions can be violated $Q_W(\eta^m) > 0$.
- Prefer markup choice η^m that leads to smaller violations (in GMM distance).
  - Calculating $\widehat{\sigma}$ is often complicated $\rightarrow$ bootstrap.
  - Duarte, Magnolfi, Sullivan (2020) show RV outperforms Cox-type tests in simulation.

Theoretical Result (more in paper)

Proposition 1a: Standard GMM assumptions, fix W and $h(x, \mathrm{w})$

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

- $\mathbb{E}[Z'\,\omega^1]$ 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 $A(\boldsymbol{z_t}) = \mathbb{E}[\Delta\eta_{jt}|\boldsymbol{z_t}]$.

Procedure

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

1. Estimate $\widehat{\omega}^m_{jt}$ as residual from (no IV necessary):

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

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

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

3. Compute the (scalar) moment violation: $\tilde{Q}(\eta^m) = \left(\frac{1}{N}\sum_{j,t} \hat{g}(\boldsymbol{z_t}) \cdot \widehat{\omega}^m_{jt}\right)^2$
4. Compare $T = \frac{\sqrt{n}}{\sigma}(\tilde{Q}(\eta^1) - \tilde{Q}(\eta^2))$ to critical values of normal after estimating $\widehat{\sigma}$ using bootstrap following Rivers and Vuong.

All regressions via random forest. (Note: different $\eta^1, \eta^2 \rightarrow$ different $A(\boldsymbol{z_t})$).

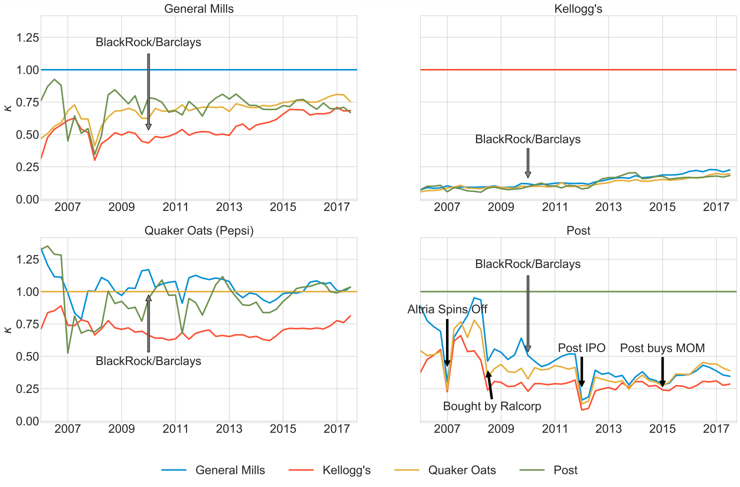
Advantages

- Fully flexible $h_s(x_{jt}, \mathrm{w}_{jt})$ (Don't specify linear, log, etc.)
- Fully flexible $\Delta\eta_{jt}^{1,2} = g(\boldsymbol{z_t}) + \zeta_{jt}$ ($\boldsymbol{z_t}$ 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 bootsrap).

Demand Estimation

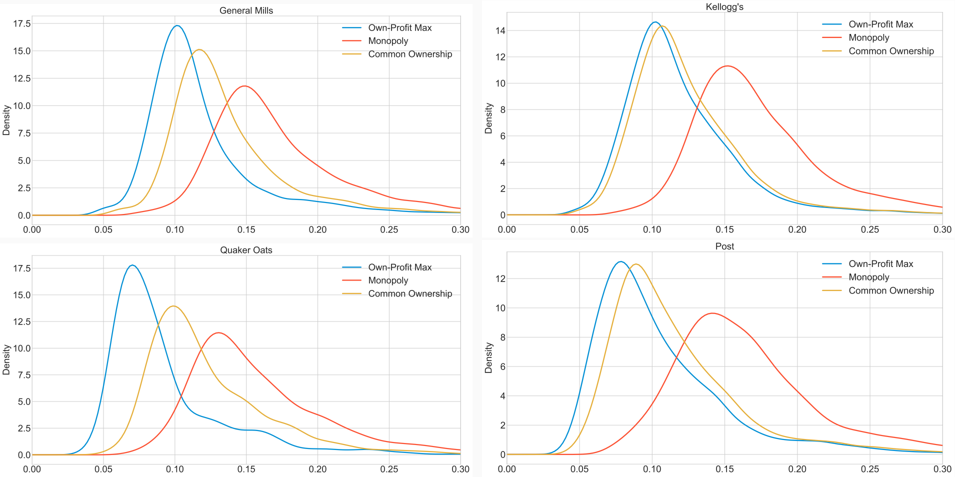
- Discrete choice demand system based on BLP (1995), Nevo (2000/1) using Kilts Data:
    - Market: Chain-DMA-week (sampled 2/13 weeks per quarter)
    - Estimate market size from milk and egg purchases.
    - Correlated random coefficients on p_{jt} and the constant.
    - 946 product FE and 1970 chain-dma-week 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 π parameters.
- Instruments:
    - Own ingredient costs and chain specific demographic variables.
    - Quadratic Gandhi-Houde differentiation instruments
    - Calculate feasible approximation to optimal instruments (18): $\mathbb{E}\left[\frac{\partial \xi_{jt}}{\partial \theta} | \boldsymbol{z_t}\right]$.
- Estimation in `PyBLP` (Conlon and Gortmaker 2020).

Why RTE Cereal?



$C4 \simeq 85\%$ domestic, public firms and good ownership variation.

Markups in dollars (Q4 2016)



Counterfactual Mergers

Firm	GM-KEL	Monopoly	κ^{CO}
General Mills	4.69	9.42	3.97
Kellogg's	5.13	9.30	5.34
Quaker Oats	-0.37	14.87	7.75
Post	-0.15	12.76	7.06
Price Index	3.32	10.25	5.42

NB: Computed using marginal costs as predicted by own-profit maximization.

Main Results: Assuming Linearity

	Others' Costs	Demographics	BLP Inst.	Dmd. Opt. Inst.
Own Profit Max vs.	Panel 1: $A(z_t) = z_t$, linear $h_s(\cdot)$; $W = (Z'Z)^{-1}$			
Common Ownership	-2.4732	-0.0079	-1.2333	-4.9099
Common Ownership (MA)	-2.5918	0.0070	-1.2105	-4.9215
Common Ownership (Lag)	-2.5208	0.0075	-1.2125	-4.9351
Perfect Competition	0.8611	-2.3033	-3.1652	-10.9229
Monopolist	-2.4166	-0.8783	-3.5162	-6.0048
Own Profit Max vs.	Panel 2: $A(z_t) = \mathbb{E}[\Delta\eta^{12}\|z_t]$, 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 $\mathbb{E}[\Delta\eta_{jt}|\boldsymbol{z_t}]$ is equivalent to a different choice of W.

Main Results: Our (Semiparametric) Test

	Others' Costs	Demographics	BLP Inst.	Dmd. Opt. Inst.
Own Profit Max vs.	Panel 3: $A(z_t) = \mathbb{E}[\Delta\eta^{12}\|z_t]$, random forest $h_s(\cdot)$ and $g(\cdot)$			
Common Ownership	-4.8893	-5.4460	-5.4412	-5.9585
Common Ownership (MA)	-5.4345	-6.1348	-5.8757	-6.4357
Common Ownership (Lag)	-5.1770	-5.9221	-5.7041	-6.2255
Perfect Competition	-7.7749	-8.7051	-8.9758	-10.0654
Monopolist	-5.2711	-6.7789	-5.9158	-6.5933

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 $h_s(\cdot), g(\cdot)$.
    - $h_s(x_{jt}, \mathrm{w}_{jt})$ 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 κ represent the weight a firm places on competitors and τ the internalization of those weights.

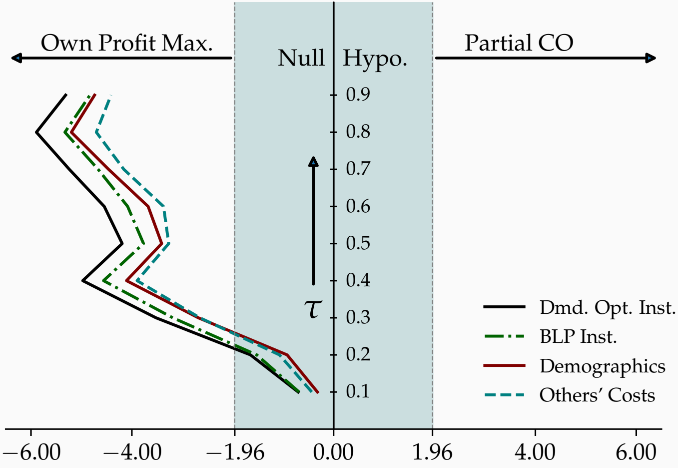
$$arg \max_{p_j : j \in \mathcal{J}_f} \sum_{j \in \mathcal{J}_f} (p_j - mc_j) \cdot s_j(\boldsymbol{p}) + \sum_{g \neq f} \tau \cdot \kappa_{fg} \sum_{j \in \mathcal{J}_g} (p_k - mc_k) \cdot s_k(\boldsymbol{p})$$

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 Testing Results



Conclusion

Our testing procedure has advantages over previous approaches:

- Amounts to two prediction exercises.
- We use the model itself to form $A(\boldsymbol{z_t})$.
- 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 η_{jt} is testable subject to relevance $\mathbb{E}[\boldsymbol{z_t'}\Delta\eta_{jt}]$.

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