Advances in Testing for the Nature of Competition

Michael Sinkinson
Yale University

Based on a research agenda with Matthew Backus (Berkeley) and Christopher Conlon (NYU)

October 30, 2023

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?

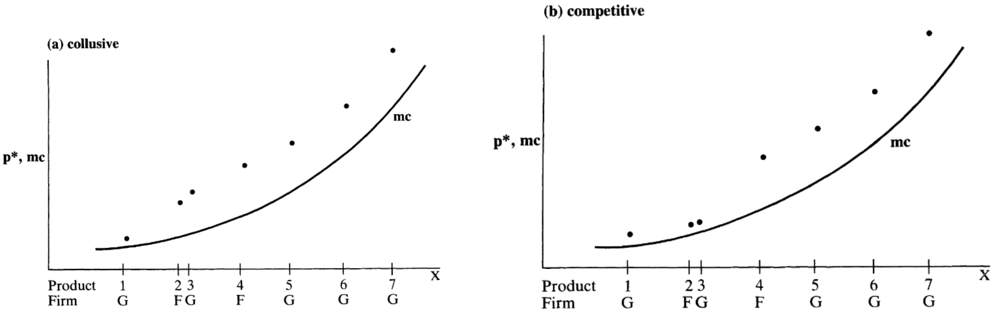
- Early work: Porter (1983), Bresnahan (1982,1987)
- Subsequent work defined the “menu” approach: Nevo (1998, 2001), Villas-Boas (2007)
- Recent revival of “internalization” parameters: Miller and Weinberg (2017), Crawford, Lee, Whinston, and Yurukoglu (2017), Pakes (2017)
- Parallel work by: Duarte, 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), Roussille and Scuderi (2022), Calder-Wang and Kim (2023), Adão Costinot and Donaldson (2023), 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.
  - A correlated shock to mc could look a lot like collusion.
- 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. (Berry Haile 2014).

A famous plot (Bresnahan 87)



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

Villas Boas (2007)

TABLE 10

p-Values for pairwise non-nested comparisons

H_0 model	Alternative models 1	2	3·1	3·2	4	5	6
1: Simple linear pricing	–	0·50	0·00	0·50	0·24	0·00	0·50
2: Hybrid	0·00	–	0·50	0·50	0·12	0·00	0·50
3.1: Zero wholesale margin	0·41	0·29	–	0·05	0·50	0·39	0·07
3.2: Zero retail margin	0·39	0·40	0·05	–	0·50	0·39	0·17
4: Wholesale collusion	0·49	0·48	0·50	0·50	–	0·48	0·50
5: Retail collusion	0·00	0·00	0·50	0·50	0·22	–	0·50
6: Monopolist	0·34	0·35	0·17	0·31	0·48	0·34	–
Chain size weighted							
1: Simple linear pricing	–	0·08	0·01	0·06	0·08	0·00	0·00
2: Hybrid	0·17	–	0·15	0·22	0·00	0·06	0·14
3.1: Zero wholesale margin	0·08	0·15	–	0·11	0·15	0·12	0·00
3.2: Zero retail margin	0·01	0·07	0·00	–	0·09	0·01	0·00
4: Wholesale collusion	0·00	0·05	0·04	0·09	–	0·00	0·02
5: Retail collusion	0·00	0·02	0·03	0·11	0·02	–	0·00
6: Monopolist	0·10	0·20	0·00	0·15	0·20	0·14	–

Notes: *p*-Values reported from non-nested, Cox-type (Smith, 1992) test statistics of the null model in a row being true against the specified alternative model in a column. Bottom part is a robustness check. It has the same format as above, but the non-nested comparisons are based on estimates for the case when the portion of the manufacturer's profit due to each retailer is weighted by the retailer's chain size.

Source: My calculations.

Conduct Testing in Pictures (Berry Haile 2014)

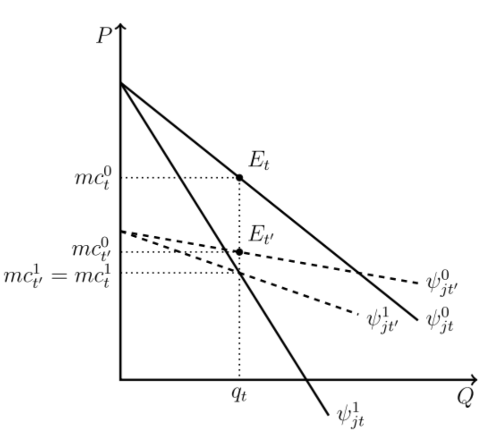


Figure 2(ab) from Berry and Haile (2014), Example 1.

Backus, Conlon and Sinkinson: Testing Common Ownership in RTE Cereal

We are interested in testing the "common ownership" model, which posits that firms maximize their investors' portfolio values, against standard models of competition.

- Common ownership predicts that firm f places a profit weight κ_{fg} on firm g. Profit weights are computed based on overlapping ownership (and some assumptions about how the firm aggregates investor preferences).
- Previous work (BCS 2021) showed that the profit weights implied by current ownership patters have grown significantly and are now over 0.7 for a typical pair of S&P 500 firms.
- In a differentiated product Bertrand world, profit weights imply an intermediate model of competition between own-profit maximization and full collusion.
- Our goal: to come up with a test that is able to distinguish between models of competition that are "close".
- RTE cereal: great setting as significant variation in implied profit weights.

Setup: Notation and Utility

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

- Markets t
- Products j
- Data $\chi_t = \{(\mathrm{x}_{jt}, \mathrm{v}_{jt}, \mathrm{w}_{jt}) \text{ for all } j \in \mathcal{J}_t\}$.
- 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: Multiproduct Bertrand Example

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.
  - We will estimate demand with rich scanner data + panelist data for demographics.
  - Estimated model allows for very flexible substitution patterns among products.
  - We treat demand as then known and use different models of $\mathcal{H}(\kappa)$ to back out implied markups under alternative models.

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))

Our Motivating Idea: 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:

- 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).

Overview of the Test (Rivers and Vuong, 2002)

We are working in a non-nested model comparison framework

- Assume demand is known, so η^1 and η^2 are also known.
- "Criterion function" matches the scalar moment

$$Q^m \equiv \mathbb{E}[\omega_{jt} \cdot \mathbb{E}[\eta^1_{jt} - \eta^2_{jt}|z_t]]^2$$

Overview of the Test (Rivers and Vuong, 2002)

We are working in a non-nested model comparison framework

- Assume demand is known, so η^1 and η^2 are also known.
- "Criterion function" matches the scalar moment

$$Q^m \equiv \mathbb{E}[\omega_{jt} \cdot \mathbb{E}[\eta_{jt}^1 - \eta_{jt}^2 | z_t]]^2$$

- $H_0 : Q_1 - Q_2 = 0$ vs $H_a : Q_1 > Q_2$ OR $Q_1 < Q_2$

Overview of the Test (Rivers and Vuong, 2002)

We are working in a non-nested model comparison framework

- Assume demand is known, so η^1 and η^2 are also known.
- "Criterion function" matches the scalar moment

$$Q^m \equiv \mathbb{E}[\omega_{jt} \cdot \mathbb{E}[\eta^1_{jt} - \eta^2_{jt}|z_t]]^2$$

- $H_0 : Q_1 - Q_2 = 0$ vs $H_a : Q_1 > Q_2$ OR $Q_1 < Q_2$
- RV show that

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

- Getting the SD of the difference is hard $\rightarrow$ bootstrap

Possible Exclusion Restrictions

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

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

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

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).
- Costs 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:

$$\sigma_j^{-1}(\mathcal{S}_t, \mathbf{p_t}, \mathrm{y_t}; \widetilde{\theta}_2) = h_d(\mathrm{x}_{jt}, \mathrm{v}_{jt}; \theta_1) - \alpha\, p_{jt} + \lambda \log(\mathrm{ad}_{jt}) + \xi_{jt}$$
$$p_{jt} - \eta_{jt}(\mathcal{S}_t, \mathbf{p_t}; \theta_2, \mathcal{H}_t(\kappa)) = h_s(\mathrm{x}_{jt}, \mathrm{w}_{jt}, q_{jt}; \theta_3) + \omega_{jt}$$

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).
- Costs 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)

Algorithm

(a) Estimate the marginal cost function under models 1 and 2 to obtain residuals $\widehat{\omega_{jt}^1}$ and $\widehat{\omega_{jt}^2}$:

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

(b) Estimate the "first stage" regression, and compute the fitted values $\widehat{\Delta\eta}_{jt}^{1,2} = \hat{g}(\mathbf{z_{jt}})$ of:

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

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

$$\tilde{Q}(\eta^m) = \left(n^{-1} \sum_{j,t} \frac{\hat{\omega}_{jt}^m}{\hat{\sigma}_{\omega}^m} \cdot \hat{g}(\mathbf{z_{jt}}) \right)^2. \tag{1}$$

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

(e) Compute the test statistic

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

Note: Steps (a) and (b) can be done in any order via non-parametric regression. Our preferred method is random forest regression which scales well as n becomes large and is well-suited to capturing nonlinear relationships.

Limitations

Not everything is testable:

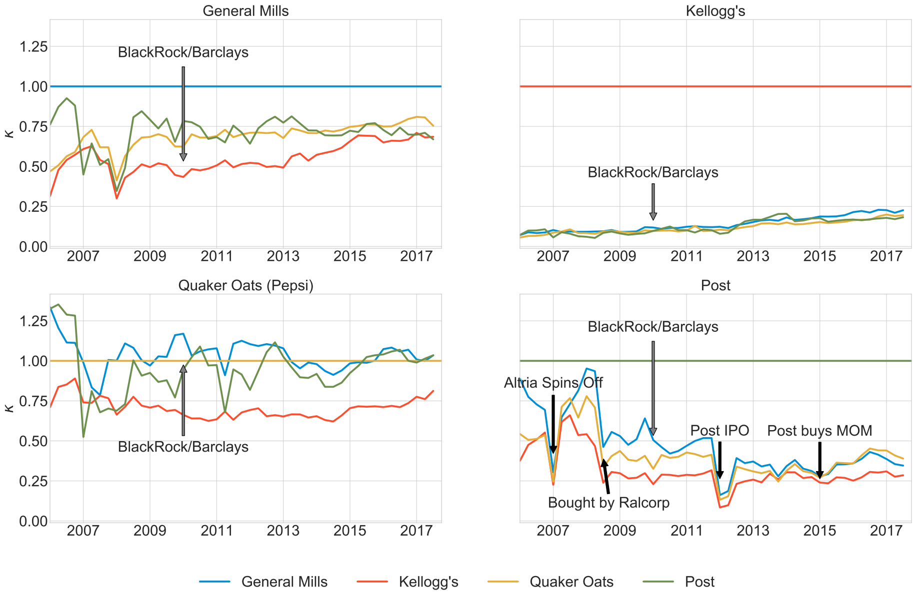
- If $\Delta\eta_{jt}$ cannot be explained by z^s_{jt} beyond contents of $(\mathrm{x}_j, \mathrm{w}_j)$ this doesn't work
- Flexible demand models are required to generate cross sectional variation in markups
- Beware of "accidental" exclusion restrictions.

Cereal Data

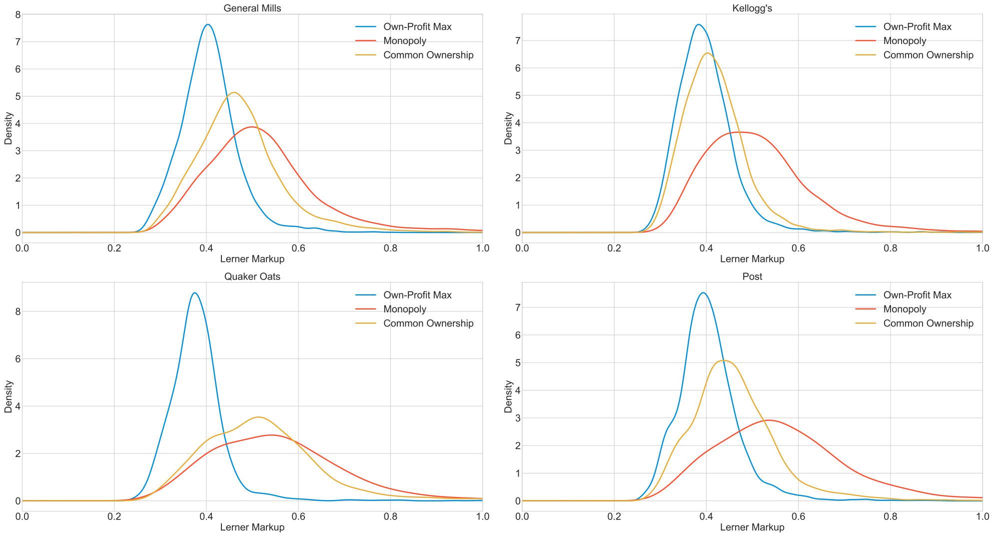
Main Dataset is NielsenIQ (from Kilts) from 2007-2017

- Consolidate to `dma-chain-week`.
- Keep largest chains who price at chain level
  - Select based on # observations from panelist data.
- Consolidate `upc` $\rightarrow$ `brand` (Honey Nut Cheerios) from multiple package sizes and box designs.
  - Divide revenue by servings
  - Maintain the fiction that households purchase servings.

Cereal Data: Variation in κ



Predicted Markups (Q4 2016)



Counterfactual Price Increases

	GM-KEL	GM-QKR	GM-POST	KEL-QKR	KEL-POST	QKR-POST	Monopoly	κ^{CO}
General Mills	8.08	1.69	3.72	0.03	0.05	0.00	14.52	5.51
Kellogg	7.79	0.01	0.04	1.68	3.77	0.00	14.25	7.89
Quaker Oats	-0.12	8.78	-0.10	9.02	-0.09	4.39	23.44	10.85
Post	0.02	-0.02	8.32	-0.01	8.82	1.91	20.27	9.61
Price Index	5.25	1.16	2.46	1.21	2.63	0.59	14.23	6.67

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

Main Results

	Others' Costs	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	-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: $A(\mathbf{z}_t) = \mathbb{E}[\Delta\eta^{12}\vert\mathbf{z}_t]$, 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
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	-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

An Internalization Parameter

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

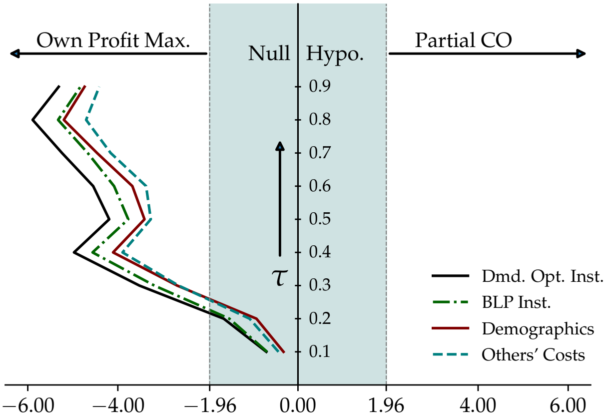
$$arg \max_{p_j \,:\, j \in \mathcal{J}_f} \sum_{j \in \mathcal{J}_f} (p_j - mc_j) \cdot s_j(\mathbf{p}) + \sum_{g \neq f} \tau \kappa_{fg} \sum_{j \in \mathcal{J}_g} (p_k - mc_k) \cdot s_k(\mathbf{p})$$

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



Stepping Back

- In order to evaluate the common ownership hypothesis, we developed a conduct testing procedure building on the identification results of Berry and Haile (2014)

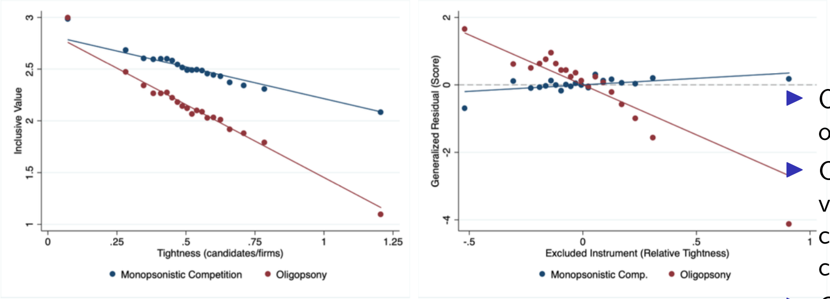
Stepping Back

- In order to evaluate the common ownership hypothesis, we developed a conduct testing procedure building on the identification results of Berry and Haile (2014)
- Our approach can test any pair of models such that
  1. Are fully specified, i.e. predict markups.
  2. Yield distinct markups ($\Delta\eta \neq 0$).
  3. We have instruments that are relevant to $\Delta\eta$, excluded from the supply function $h_s(\cdot)$, and mean independent of ω^0.

  - Cartels: collusive versus oligopoly pricing?
  - Vertical contracts: DM or manufacturer pricing?
  - Labor: monopsony versus perfectly competitive labor markets?
  - Behavioral: suboptimal versus rational pricing rules?
  - Next, some recent examples

Scuderi JMP 2022: Models of Labor Supply

Figure 5: Vuong Test



(a) First Stage   **(b)** Visualizing the Vuong Test

Note: Panel (a) plots the "first stage" relationship between the model-implied inclusive values Λ and Λ_i^{-j} and the instrumental variable t_{ij}, conditional on firm covariates z_j and candidate covariates x_i and two-week period dummies. Panel (b) plots the relationship between generalized residuals and the excluded instrument for the non-predictive monopsonistic competition and oligopsony models. Under proper specification, the correlation of the generalized residuals and the excluded instrument should be zero (the dashed line). The larger the deviation from zero, the greater the degree of mis-specification of the model.

- Offered wages for online job platform
- Compares Monopsony vs. oligopolistic competition vs. perfect comp.
- Compares tailored offers vs. not (price discrimination).

Scuderi JMP 2022: Models of Labor Supply

- Firms ignore competitors (Monopsony)
- Firms offer wages independent of candidate characteristics (experience, demographics).
- Firms are definitely NOT paying MPL.

Table 4: Non-Nested Model Comparison Tests (Rivers and Vuong, 2002

Model	(1) Monopsonistic Comp.	(2)	(3) Oligopsony	(4)
	Not Predictive	Type Predictive	Not Predictive	Type Prec
Perfect Competition	-54.84	-54.40	-39.92	-39.9
Monopsonistic, Not Predictive	–	7.83	3.98	2.69
Monopsonistic, Type Predictive		–	2.77	1.54
Oligopsony, Not Predictive			–	-3.67
Oligopsony, Type Predictive				–

Note: This table reports test statistics from the Rivers and Vuong (2002) non-nested model comparison procedure. values imply the row model is preferred to the column model. Under the null of model equivalence, the test stati asymptotically normal with mean zero and unit variance.

specification. In the figure, the generalized residuals for the monopsonsitic cor tion alternative are closely aligned with the x-axis, while the generalized residu the oligopsony alternative are strongly negatively related to tightness.

Our tests therefore suggest that models of firm behavior in which firms both i strategic interactions in wage setting and do not tailor wage offers to candida the basis of predictable preference variation are closer approximations to firms bidding behavior on the platform than are models in which firms act strateg and tailor offers. In Appendix F, we report the testing results using the or Vuong (1989) likelihood comparison test, which yield qualitatively identical comparisons. In the following analysis, we adopt the not-predictive monopso competition model as our preferred model of conduct.

Starc Wollmann: Generic Pharma Cartel + Entry

Do Cartels encourage entry with high prices?

> *in 2013, Teva Pharmaceuticals, the largest generic firm, hired NP, a marketing executive with especially strong industry relationships, and tasked her with "price increase implementation."1 Over an 18-month period, industry participants exchanged thousands of calls and texts—alongside countless LinkedIn, Facebook, and WhatsApp messages and face-to-face conversations—with contacts at rival firms to coordinate the increases (Complaint, page 322).2 Following this period, prescription drug expenditures by governments, private insurers, and individuals rose sharply by billions of dollars.*

Starc Wollmann: Generic Pharma Cartel

- NP organizes the cartel and prices go up
- Slightly less in large markets (which are more likely to see entry)

Figure I: *Prices rise sharply following cartel formation*

This figure plots the average log price of cartelized and uncartelized drugs on the
against calendar quarter on the x-axis. The vertical red line corresponds to the first q
of 2013—the period in which NP joined Teva. Prices are normalized to zero in that qu

Starc Wollmann: Baseline Scenario

- all firms set competitive prices in uncartelized markets
- all firms set competitive prices in cartelized markets before cartel formation;
- and after cartel formation, members set prices that maximize their joint profits while nonmembers best respond

Starc Wollmann: Testing Results

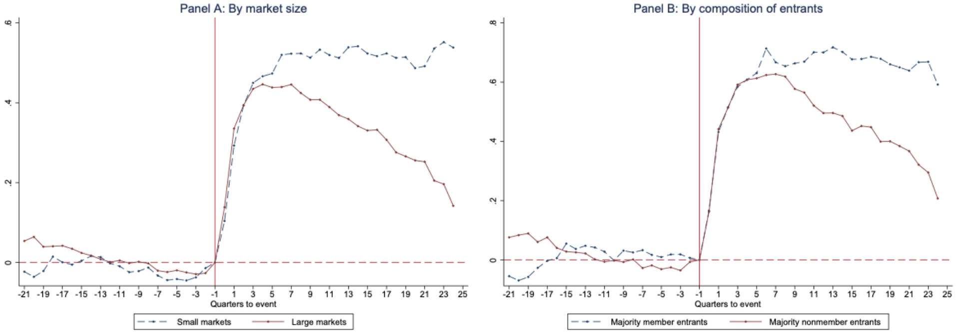


Figure III: *Prices after cartel formation*

In this figure, we plot estimates of β^τ, which are derived from equation 1, against x^τ, which represents event time in quarters. For Panel A, we proxy for market size by computing the number of prescriptions filled in the period just prior to NP joining Teva, and we distinguish between large and small cartelized markets based on how this figure relates to the median value. For Panel B, we calculate the fraction of entrants that arrive as a result of ANDAs filed after NP joins Teva and identify markets in which a majority of those firms are members and nonmembers, respectively.

Starc Wollmann: Generic Pharma Cartel

Table IV: *Results of conduct tests*

	$\tilde{Q} \times 100$ Baseline	Alternative	Test statistic
Test A. Model vs. competition (pre-investigation)	.16247	.20688	-3.03
Test B. Model vs. competition (post-investigation)	.20286	.34235	-4.63
Test C. Model vs. nonmembers comply	1.3979	4.6533	-5.57
Test D. Model vs. member entrants do not comply	.00008	.00008	-3.06

This table reports the results of the testing procedure proposed by Backus et al. (2021) for pairwise described in the text. The test statistic is distributed standard normal. The standard error of t between $\tilde{Q}_1$ and $\tilde{Q}_2$ is obtained via bootstrapping.

Wrapping Up

Takeaways:

- Equilibrium markups are a nonlinear function of everything in the model. Using the model to get that nonlinearity right makes for a more powerful conduct test.
- In RTE cereal, we see strong evidence in favor of own-profit maximization rather than common ownership pricing.
- Discussion
  - We reject evidence of CO short-run price comepetition in RTE Cereal.
  - Can't reject other mechanisms (CEO's living the quiet life)
  - Can't explain stock market pricing anomalies
  - In some sense CO is what we would see absent agency problems, so where are they?
  - Hopefully our testing procedure is useful in other contexts.