Market Power: Collusion and Horizontal Mergers (2)

Jean-François Houde
Cornell University & NBER

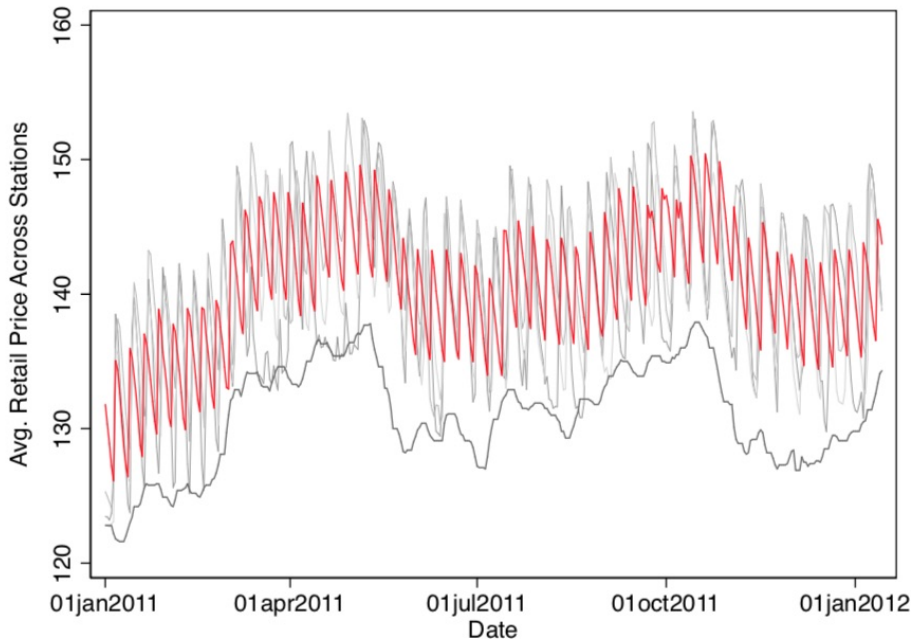
February 2, 2018

Byrne and de Roos (2017): Learning to coordinate

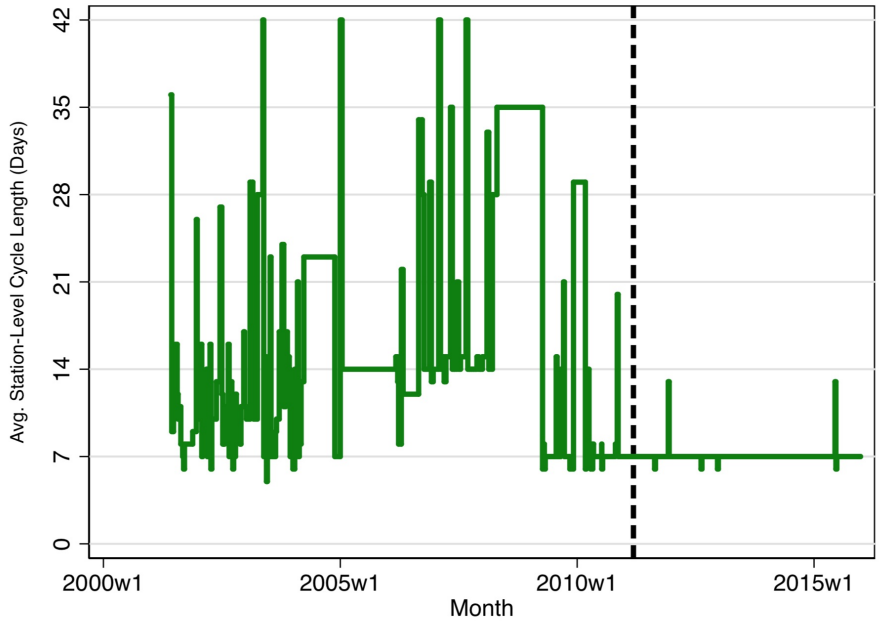
- In 2001, the Western Australia's state government introduced a price transparency regulation. What does it do?
  - Next day's price must be announced before 2pm
  - This price is valid for 24hrs (from 6am)
- First-order effect of the regulation: Mis-pricing is *very* costly (24hrs delays)!
- **Market:** Perth's gasoline retail market
  - Integrated retailers: BP, Caltex, Mobil, and Shell
  - Grocery store chains: Coles and Woolworths

Time-series: Price Cycles

Prices and costs in 2011
across 4 australian cities



The collapse of
cycle lenghts in Perth



- **Note:** Red = Perth (other cities are not price regulated).
- Prior to 2010, prices and profits margins were more unstable in Perth than in other cities.

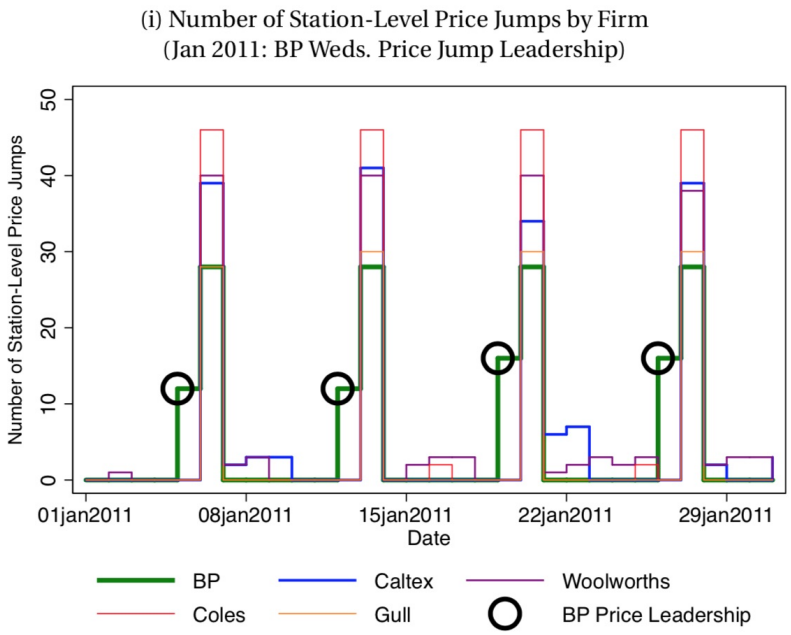
Key features of the price adjustment strategy

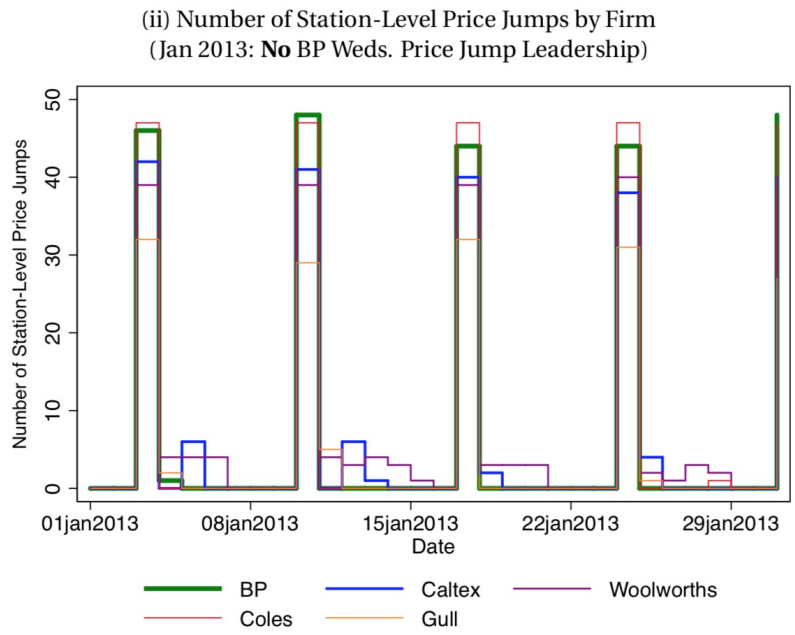
- Starting in 2010, firms coordinated on the following pricing rule:
  1. Thursday price *jump*
  2. Daily price cuts by 2 cpl
  3. Cycle length = 7 days
- The size of price increase is determined by expected margins at the bottom of the cycle:

$$p_{t+7} = p_t - 12 \geq E(c_{t+7})$$

- **Implications for profits:** Prices change every day, but profit margins are determined by the Thursday price.
- **Coordination problem:** How to determine the Thursday price given the constraints imposed by the price-posting regulation?

Pricing cycles with and without leadership

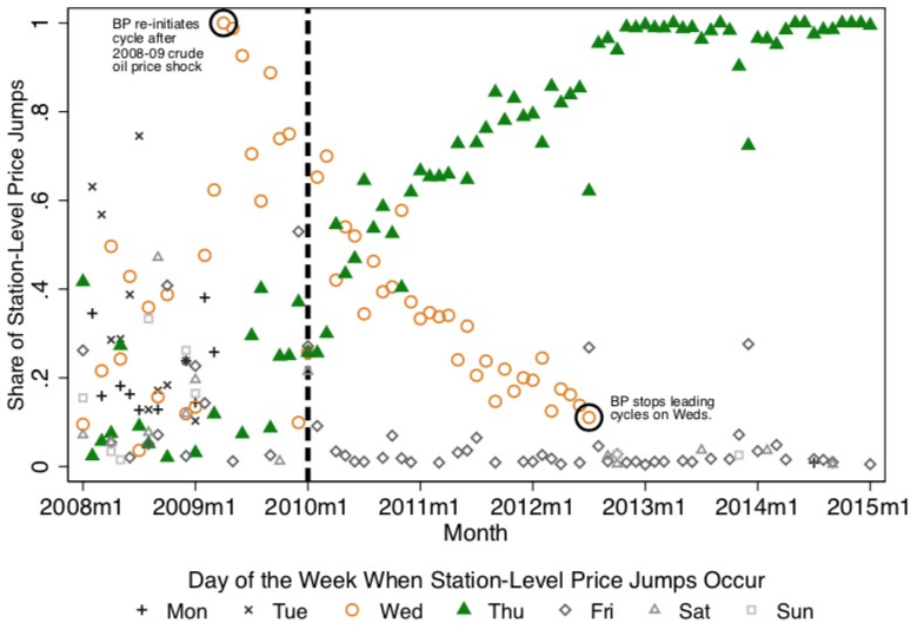




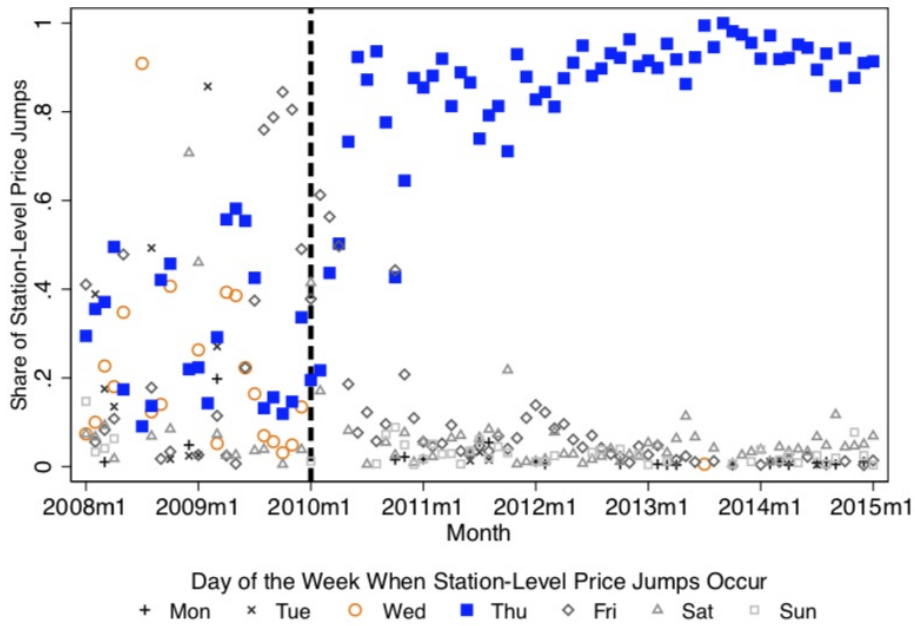
- The strategy was implemented gradually, with BP taking a **price leadership** role
  - **Sequential-move period:** BP (largest network) "jumps" first on Wednesday (give up market share), and rivals price match on Thursday
  - **Simultaneous-move period:** All major chains "jump" simultaneously on Thursday to re-start the cycle

Transition period (1): BP gradually decreasing its share of stations moving first on Wednesday

(i) Share of **BP** Station-Level Price Jumps by Day of the Week and Month

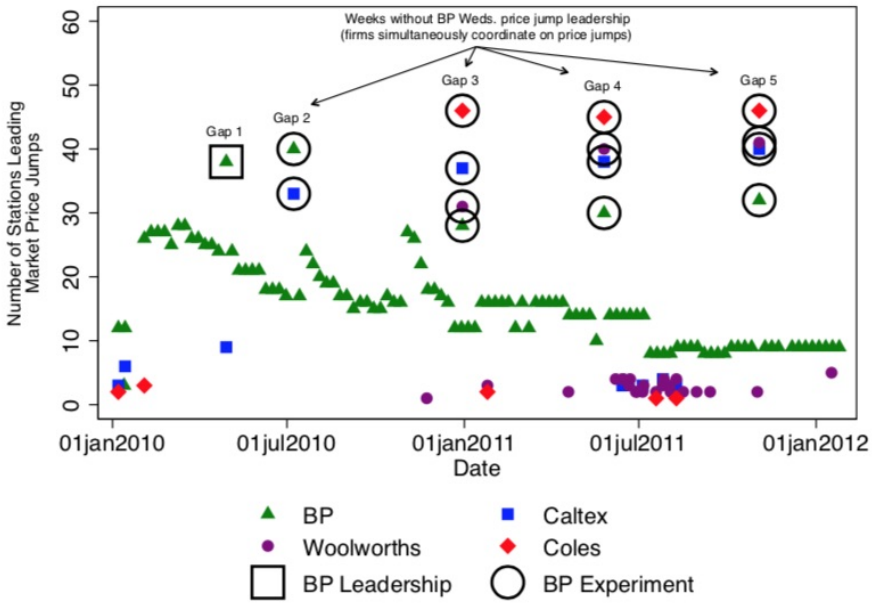


(ii) Share of **Caltex** Station-Level Price Jumps by Day of the Week and Month

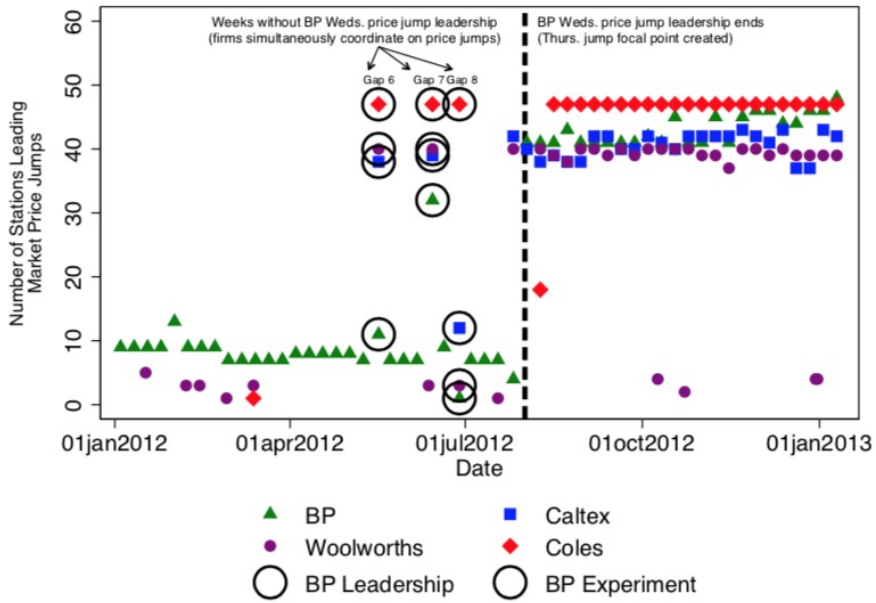


Transition period (2): BP periodically experiments with Thursday price jumps (signals to rivals)

(i) Number of Stations Leading Market Price Jumps by Firm
Jan 2009 – Jan 2012

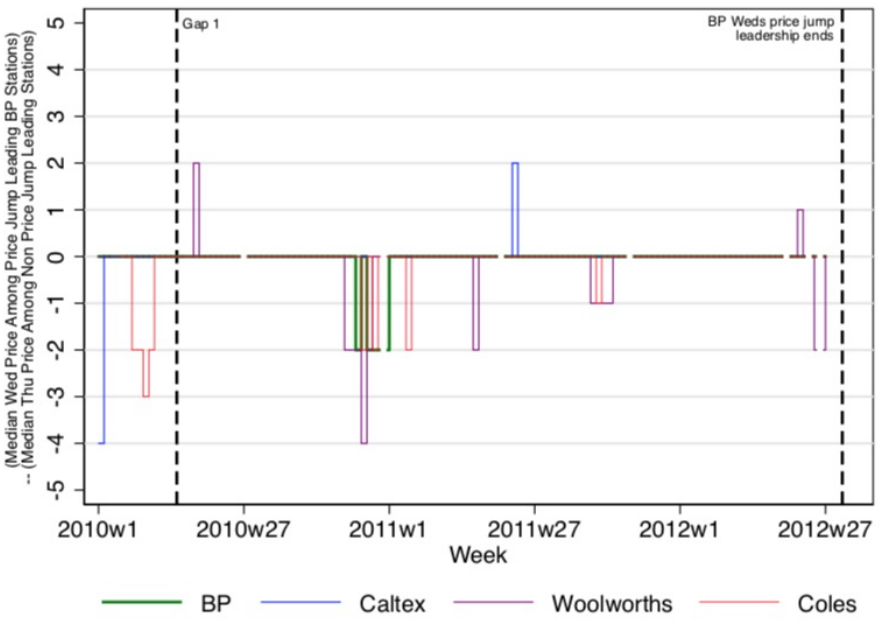


(ii) Number of Stations Leading Market Price Jumps by Firm
Jan 2012 – Jan 2013

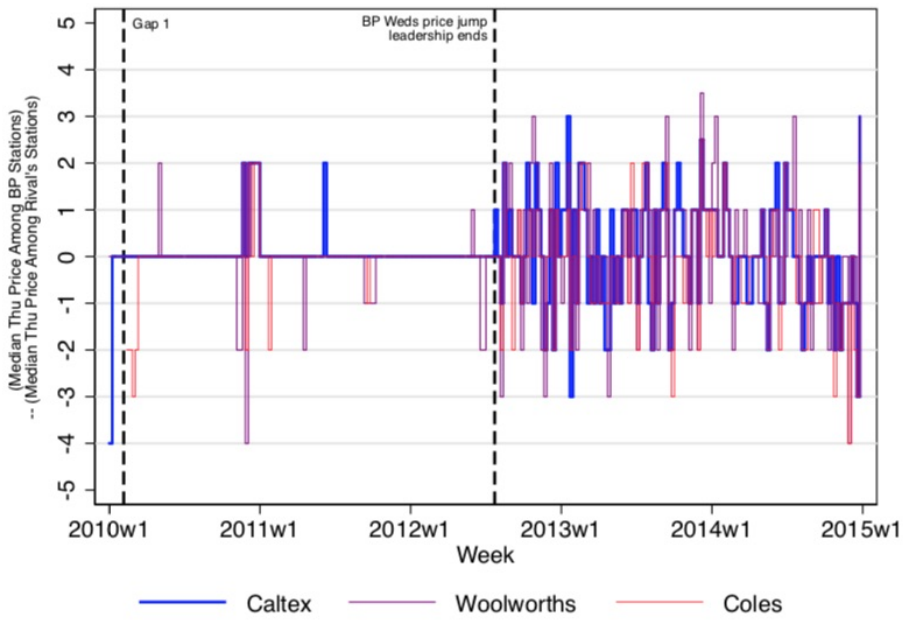


Simultaneous equilibrium: Coordination Problems on Thursday Prices

(i) Difference Between Median **Wed.** Price Among Price Jump Leading BP Stations and Median **Thu.** Price Among Non-Leading Stations



(ii) Difference Between Median **Thu.** Price Among All BP Stations and Median **Thu.** Price Among Rival Stations

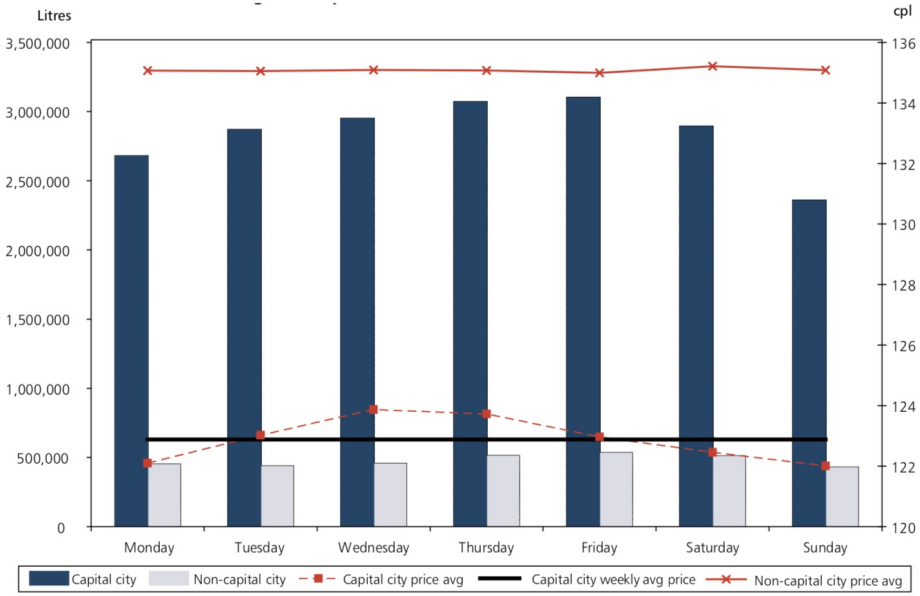


- **Takeaway:** The role of the first-mover is important to provide stability and information on the collusive price.

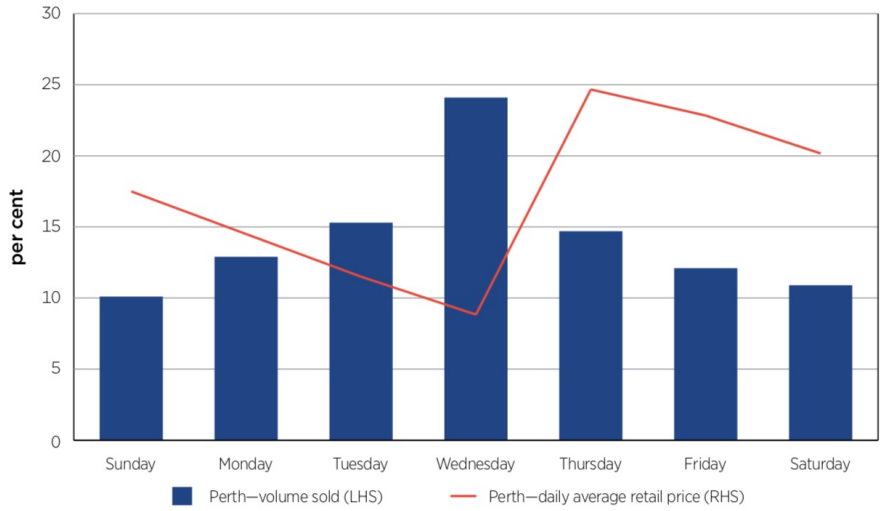
Implication for consumers?

Demand and Pricing Cycles in Perth

Period: 2003-2004



Period: 2013-2014



- Within-week cycle $\approx$ Price discrimination device between informed and uninformed consumers (e.g. Airline tickets)
- This also implies that it is very costly to move first in this market (25% of sales on Wednesday!)

Conduct in differentiated product markets

- Consider a market with n differentiated products:

$$s_j = D_j(p_j, p_{-j})$$

  where s_j is the market-share of j (including outside good).
- Profit maximization leads to the following FOC:

$$D_j(p_j, p_{-j}) + \sum_k \theta_{jk} \frac{\partial D_k(p_k, p_{-k})}{\partial p_j}(p_k - c_k) = 0$$

- **Conduct**: θ_{jk} measures the portion of profits generated from product k that is *internalized* by product j.
- **Examples:** Bertrand-Nash equilibrium
  - Individual products firms: $\theta_{jj} = 1$, and $\theta_{jk} = 0$ for all $k \neq j$.
  - Multi products firms: $\theta_{jk} =$ for all k owned by the same firm as j (zero otherwise).
  - Monopoly: $\theta_{jk} = 1$ for all j and k.

Identifying Industry Conduct

- The previous FOC leads to the following supply relation (in matrix form):

$$p = c + \Sigma s$$

- Markup are determined by: $\Sigma = -\left[\Delta \cdot \Theta\right]^{-1}$, a $n \times n$ matrix with (j, k) element,

$$\left[\Sigma^{-1}\right]_{jk} = \theta_{jk} \cdot \frac{\partial D_k}{\partial p_j}$$

- Clearly, even if we assume a constant marginal-cost function, Θ contains too many parameters to be identified in practice (Nevo, 1998).
- The literature has therefore split into two branches:
  - **Structural approach:** Which conduct assumption best describes the data? And more recently, estimation of a parametrized version of the conduct matrix Θ.
  - **Reduced-form approach:** What is the impact of "exogenous" conduct changes on prices?

Early Example: The 1955 price war in the US auto market

Source: Bresnahan, JIE, 1987

TABLE I

Year	(1) Auto Production[a]	(2) Real Auto Price-CPI[b]	(3) % Change Auto Price-Cagan[c]	(4) Auto Sales[d]	(5) Auto Quantity Index[e]
1953	6.13	1.01	NA	14.5	86.8
1954	5.51	0.99	NA	13.9	84.9
1955	7.94	0.95	−2.5	18.4	117.2
1956	5.80	0.97	6.3	15.7	97.9
1957	6.12	0.98	6.1	16.2	100.0

Notes: [a] Millions of units over the model year. [Source: *Automotive News.*]
[b] (CPI New automobile component)/CPI. [Source: *Handbook of Labor Statistics.*]
[c] Adjusted for quality change. [See Cagan (1971), especially pp. 232–3.]
[d] Auto output in constant dollars, *QIV* of previous year through *QIII* of named year, in billions of 1957 dollars. [Source: *National Income and Product Accounts.*]
[e] (4)/(2), normalized so 1957 = 100.

- **Research question:** Is the 1955 increase in production explained by a deviation from collusion?

Quality Ladder Model

- J vertical differentiated products: $x_J > x_{J-1} > ... > x_2 > x_1 > x_0$.
- Indirect utility given marginal utility for quality v_i:

$$U(x_j, v_i, Y) = \begin{cases} v_i x_j + Y - P_j & \text{If } j \neq 0 \\ v_i x_0 + Y - E & \text{If } j = 0 \end{cases}$$

- Distribution assumption: $v_i \sim U(0, V_{max})$ with density δ.
- Market shares:

$$s_j = D_j(p_j, p_{-j}) = \begin{cases} \frac{1}{\delta}\left(V_{\max} - \frac{P_J - P_{J-1}}{x_J - x_{J-1}}\right) & \text{If } j = J, \\ \frac{1}{\delta}\left(\frac{P_{j+1} - P_j}{x_{j+1} - x_j} - \frac{P_j - P_{j-1}}{x_j - x_{j-1}}\right) & \text{If } 1 \geq j < J, \\ \frac{1}{\delta}\left(\frac{P_1 - E}{x_1 - x_0}\right) & \text{If } j = 0. \end{cases}$$

- Constant marginal cost function: $mc(x)$ and $mc'(x) > 0$.

Firm conduct assumptions

- Individual product Bertrand-Nash:

$$s_i + (P_i - mc(x_i))\frac{\partial D_i}{\partial P_i} = 0$$

- Multi-product Bertrand-Nash:

$$s_i + \sum_{j=i-1}^{j=i+1} \theta_{ij}(P_j - mc(x_j))\frac{\partial D_j}{\partial P_i} = 0$$

  Where $\theta_{ij} = 1$ if products i and j are produced by the same firm

- Collusion:

$$s_i + \sum_{j=i-1}^{j=i+1} (P_j - mc(x_j))\frac{\partial D_j}{\partial P_i} = 0$$

Identification in Figure

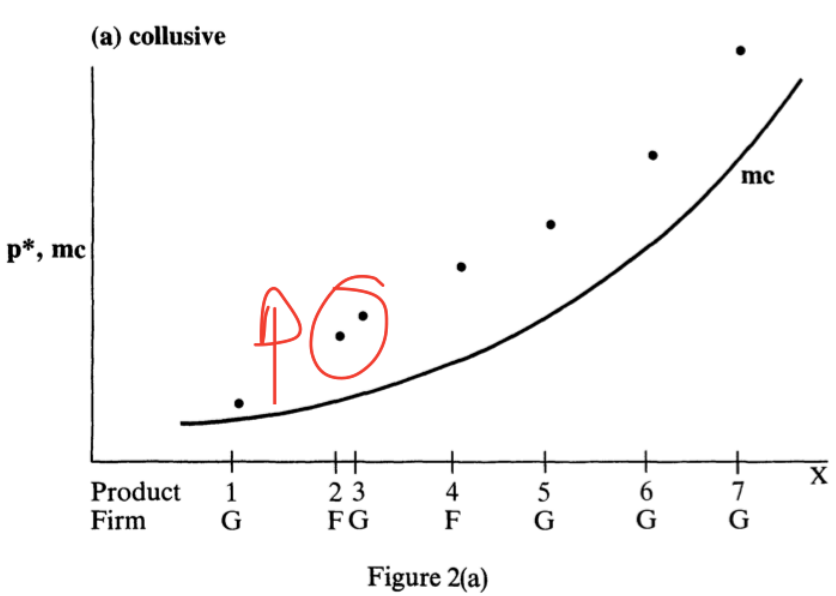


Figure 2(a)

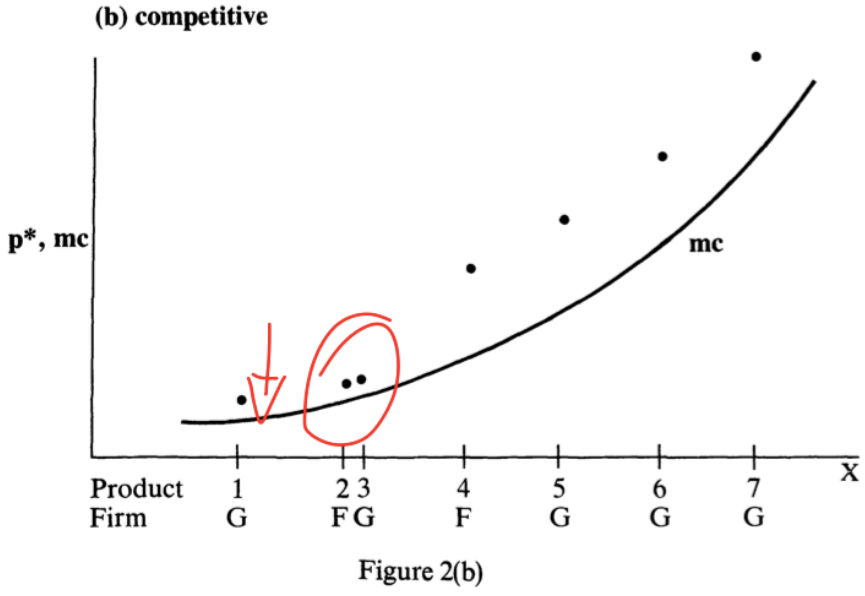


Figure 2(b)

- If the marginal cost function is monotonically increasing in quality, prices should be increasing in quality independently of the presence of closed substitutes (under collusion).
- If firms are competing, prices should also be function of the presence of close substitutes.

Testing for Collusion in Practice

- Empirical model:
  - Deterministic quality and marginal cost functions:

$$x_i = \sqrt{\beta_0 + \beta' z_i}$$

$$mc(x) = \mu exp(x)$$

    where z_i is vector of car characteristics.
  - Measurement errors:

$$p_i = p_i^*(x|H,\theta) + \epsilon_i^p$$

$$q_i = q_i^*(x|H,\theta) + \epsilon_i^q$$

    where $(\epsilon_i^p, \epsilon_i^q) \sim N(0, \Sigma)$.
  - Likelihood function:

$$L(P, Q|Z, \theta, H) = \sum_i \ln f(p_i - p_i^*, q_i - q_i^*|\Sigma)$$

- **Non-nested hypothesis test:** Likelihood ratio of H_0 and H_1, evaluated under H_0 (Cox statistic).

Result: Bertrand-Nash cannot be rejected in 1955

TABLE III
COX TEST STATISTICS

Hypotheses	*C*	*N–C*	*'p'*	*H*
a—1954				
Collusion	—	0.8951	0.9464	−1.934
Nash-Competition	−2.325	—	−0.8878	−2.819
"Products"	−3.978	3.029	—	−1.604
Hedonic	−12.37	−10.94	−13.02	—
b—1955				
Collusion	—	−10.36	−9.884	−13.36
Nash-Competition	−1.594	—	1.260	0.6341
"Products"	−0.7598	−4.379	—	−1.527
Hedonic	−3.353	−8.221	−5.950	—
c—1956				
Collusion	—	1.227	0.8263	1.629
Nash-Competition	−2.426	—	−4.586	0.8314
"Products"	−3.153	0.9951	—	4.731
Hedonic	−5.437	−9.671	−11.58	—

Incorporating Unobservable Attributes

- The Bresnahan (1987) approach ignores the possibility that **quality** and/or **cost** include unobserved attributes.
- Berry (1994) and Berry et al. (1995) (BLP) proposes an econometric model to incorporate both sources of unobserved heterogeneity **within** the model.
- Mixed-logit demand system:

$$s_j = D_j(\delta, x) = \int \frac{\exp\left(\delta_j + \mu(\eta_i, x_j)\right)}{1 + \sum_k \exp\left(\delta_k + \mu(\eta_i, x_k)\right)} dF(\eta_i)$$

  where M is the market size, and $\mu(\eta_i, x_j) = \sum_l \sigma_l \eta_i x_{j,l}$ is the consumer of type η_i's taste for the characteristics of product j.
- Notation:
  - Average value: $\delta_j = x_j \beta_x - \beta_p p_j + \xi_j$
  - Marginal cost: $\text{mc}_j = w_j \gamma_w + \gamma_s s_j + \omega_j$

Estimation

- There are broadly three approaches to quantify market power in this framework:
  1. Estimate θ and γ jointly under a conduct assumption (i.e. Θ)
  2. Estimate θ first, and test the validity of a particular conduct model among a **menu** of models.
  3. Estimate θ first, and Θ and γ together (after imposing functional form restrictions)

- Berry et al. (1995) is an example of the first category, and Nevo (2001) is an example of the second.
  - I will focus only on the second approach, and assume that we can consistently estimate demand without a conduct assumption.
- Ciliberto and Williams (Rand, 2014), Weinberg and Miller (ECMA, 2016), and Sullivan (wp, 2016) are examples of the third approach.

Menu Approach

- Identifying assumption in market t:

$$E(\omega_{jt}|x_t, w_t) = 0$$

- From this conditional independence assumption we can form non-linear moment restrictions that are valid under the **true** conduct model:

$$E\left[\left(p_{jt} - \Sigma_t^m s_t - w_{jt}\gamma_w - \gamma_s s_{jt}\right) \cdot z_{jt}\right] = 0 \quad \text{Iff } m = 0$$

$$E\left[\left(p_{jt} - \Sigma_t^m s_t - w_{jt}\gamma_w - \gamma_s s_{jt}\right) \cdot z_{jt}\right] \neq 0 \quad \text{Iff } m \neq 0$$

where $\Sigma_t^0 s_t = -\left[\Delta_t^0 \cdot \Theta_t^0\right]^{-1} s_t$ is the true vector of profit margins (unknown).

Example: Profit margins in the cereal market

Nevo, ECMA, 2001.

Accounting marginal cost and profits

DETAILED ESTIMATES OF PRODUCTION COSTS

Item	$/lb	% of Mfr Price	% of Retail Price
Manufacturer Price	2.40	100.0	80.0
Manufacturing Cost:	1.02	42.5	34.0
Grain	0.16	6.7	5.3
Other Ingredients	0.20	8.3	6.7
Packaging	0.28	11.7	9.3
Labor	0.15	6.3	5.0
Manufacturing Costs (net of capital costs)[a]	0.23	9.6	7.6
Gross Margin		57.5	46.0
Marketing Expenses:	0.90	37.5	30.0
Advertising	0.31	13.0	10.3
Consumer Promo (mfr coupons)	0.35	14.5	11.7
Trade Promo (retail in-store)	0.24	10.0	8.0
Operating Profits	0.48	20.0	16.0

[a] Capital costs were computed from ASM data.
Source: Cotterill (1996) reporting from estimates in CS First Boston Reports "Kellogg Company," New York, October 25, 1994.

Predicted profit margins

MEDIAN MARGINS[a]

	Logit (Table V column ix)	Full Model (Table VI)
Single Product Firms	33.6% (31.8%–35.6%)	35.8% (24.4%–46.4%)
Current Ownership of 25 Brands	35.8% (33.9%–38.0%)	42.2% (29.1%–55.8%)
Joint Ownership of 25 Brands	41.9% (39.7%–44.4%)	72.6% (62.2%–97.2%)
Current Ownership of All Brands	37.2% (35.2%–39.4%)	—
Monopoly/Perfect Price Collusion	54.0% (51.1%–57.3%)	—

[a] Margins are defined as $(p - mc)/p$. Presented are medians of the distribution of 27,862 (brand-city-quarter) observations. 95% confidence intervals for these medians are reported in parentheses based on the asymptotic distribution of the estimated demand coefficients. For the Logit model the computation is analytical, while for the full model the computation is based on 1,500 draws from this distribution.

- **Conclusion:** Accounting marginal-cost estimates imply that the Bertand-Nash assumption cannot be rejected against the "perfect" collusion.
- Nevo's approach is not formally testing a particular conduct assumption.

Formally testing alternative conduct assumptions

- Assume that the demand function can be consistently estimated:
  - Known substitution patterns $\Delta = \{\partial D_j/\partial p_k\}_{j=1,\ldots,J,k=1,\ldots,J}$
- Alternative conduct assumptions (i.e. different matrix Θ), lead to different supply relations (in vector form):

$$p_t = w_t\gamma_w + \gamma_s s_t + \Sigma_t s_t + \omega_t$$

  where $\Sigma_t = [\Delta_t \cdot \Theta]^{-1}$.
- Let m indexes the different conduct models.
- If z_{jt} is a vector of **valid** instruments, the validity of a particular conduct model m can be evaluated from the moment restrictions:

$$\bar{g}^m = \frac{1}{N}\sum_{jt} \hat{\omega}_{jt}^m \cdot z_{jt}$$

  where $\hat{\omega}_{jt}^m = p_{jt} - w_t\hat{\gamma}_w - \hat{\gamma}_s s_t - \Sigma_t^{m-1} s_t$ is the model residual under model m.

Example: Testing for alternative vertical pricing models

Villas-Boas, Restud, 2007.

p-Values for pairwise non-nested comparisons

	Alternative models						
H_0 model	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.

- **Conclusion:** Double marginalization models are (weakly) rejected against non-linear pricing models with zero upstream (or downstream) margins.

Example: Testing for resale-price maintenance

Bonnet et Dubois, RJE, 2010

TABLE 7 Models Estimated

Models	Retailers	Manufacturers	Vertical Interaction
Linear tariffs			
Model 1	Collusion	Nash	Stackelberg
Model 2	Nash	Collusion	Stackelberg
Model 3	Collusion	Collusion	Stackelberg
Model 4	Nash	Nash	Nash
Model 5	Nash	Nash	Stackelberg
Two-part tariffs			
Model 6	Nash	Nash	RPM ($w = \mu$)
Model 7	Nash	Nash	RPM ($p = w + c$)
Model 8	Nash	Nash	No RPM

TABLE 11 Results of the Rivers and Vuong Test (500 Bootstrap Replications)

Test Statistic $T_n = \frac{\sqrt{n}}{\hat{\sigma}_n}\left(Q_n^2(\hat{\Theta}_n^2) - Q_n^1(\hat{\Theta}_n^1)\right) \to N(0,1)$

$H_1 \backslash H_2$	2	3	4	5	6	7	8
1	2.99	6.63	2.11	2.64	−6.51	−6.94	−0.84
	(5.33)	(6.19)	(7.54)	(4.23)	(4.31)	(4.24)	(6.68)
2		−2.35	−3.12	−0.18	−5.59	−5.84	−3.68
		(4.90)	(4.52)	(4.54)	(4.22)	(4.96)	(3.98)
3			−1.60	1.91	−7.22	−7.75	−1.41
			(7.29)	(4.67)	(4.74)	(5.05)	(6.84)
4				2.85	−8.00	−8.50	−1.04
				(3.30)	(4.17)	(3.90)	(5.97)
5					−4.95	−5.24	−3.58
					(3.15)	(3.60)	(4.56)
6						−5.09	2.86
						(4.49)	(3.92)
7							4.25
							(4.59)

- **Rivers-Vuong test:** Difference between two over-identification tests is (asymptotically) normally distributed.
- **Conclusion:** The resale-price maintenance assumption cannot be rejected against all double-marginalization models.

Example: Testing for coordinated effects *across* firms

Miller and Weinberg, working paper, 2015

- While estimating Θ is very *daunting*, it is in principle feasible to test wether two firms are (partially) colluding.
- Miller and Weinberg consider the following conduct matrix:

$$\theta_{jk} = \begin{cases} 1 & \text{If } F_j = F_k \\ \kappa & \text{If } F_j \neq F_k \end{cases}$$

- κ is the degree of *coordination*
- See also Ciliberto and Williams (Rand, 2014) for a related test of the multi-market contact hypothesis.

Table 4: Supply Side Estimates

Regressor	Parameter	(i)	(ii)	(iii)	(iv)
Nonlinear Parameters					
Brewer Collusion	κ	0.330***	0.327***	0.336***	0.381***
		(0.092)	(0.096)	(0.087)	(0.090)
Retail Scaling	λ	0.068	0	0.25	1
		(0.100)	-	-	-
Linear Parameters					
MillerCoors×PostMerger	γ_1	-0.557***	-0.506***	-0.616***	-0.739**
		(0.084)	(0.111)	(0.114)	(0.319)
MillerCoors×PostMerger ×24-Pack	γ_2	-0.321***	-0.324***	-0.318***	-0.301***
		(0.061)	(0.086)	(0.065)	(0.085)
Distance	γ_3	0.167***	0.170***	0.161***	0.115***
		(0.062)	(0.070)	(0.081)	(0.151)
Distance×24-Pack	γ_4	0.018	0.007	0.029	0.088
		(0.039)	(0.051)	(0.050)	(0.086)
Derived Statistics					
Negative Marginal Costs		1.28%	0.54%	11.99%	70.94%
GMM J-Statistic		29.37	30.03	29.50	29.40

- **Application:** Competition between Miller/Coors and Anheuser-Busch Inbev (ABI).

What is a good instrument?

- The previous examples suggest that we can compare alternative conduct models by testing the validity of moment restrictions:

$$E(\omega_{jt}^0 \cdot z_{jt}) = 0$$

- In ongoing work with Amit Gandhi, we ask the two related questions:
  - What is a **relevant** restriction to identify the right conduct model?
  - How can we test alternative conduct assumptions without relying a particular FOC?
- Formally, the identification of the supply relation can be seen as a non-linear IV regression problem:

$$p_{jt} = w_{jt}\gamma_w + \gamma_s s_{jt} + \pi_{jt}^0 + \omega_{jt}^0$$

where $\pi_t^0 = \left[\Delta_t^0 \cdot \Theta_t^0\right]^{-1} s_t$ is the "true" profit margin vector in market t.

- Alternative models lead to different "marginal-cost" regressions:

$$p_{jt} - \pi_{jt}^m = w_{jt}\gamma_w + \gamma_s s_{jt} + \omega_{jt}^m$$

Two Observations

1. For a given model m, π_{jt}^m is correlated with ω_{jt}^0. OLS estimation of the following price regression is not valid:

$$p_{jt} = w_{jt}\gamma_w + \gamma_s s_{jt} + b\pi_{jt}^0 + \omega_{jt}^0$$

   That is, $\hat{b}^{ols} \neq 1$ at the true model.

2. A relevant IV is one that can explain the **dispersion** of **true** profit margins within a market (i.e. markup shifter).
   - If $\Theta^m \neq \Theta^0$ is a candidate model, the cost regression can be written as:

$$p_{jt} - \pi_{jt}^m = w_{jt}\gamma_w + \gamma_s s_{jt} + \omega_{jt}^m$$

     where the residual is: $\omega_{jt}^m = \pi_{jt}^0 - \pi_{jt}^m + \omega_{jt}^0$.
   - A good IV must **reject** the moment restrictions at the **wrong** conduct models, and **not-reject** the moment restrictions at the **true** model
   - That is, z_{jt} must be correlated with margin differences: $\pi_{jt}^0 - \pi_{jt}^m$.

Illustration: Back to Bresnahan (1987)'s picture

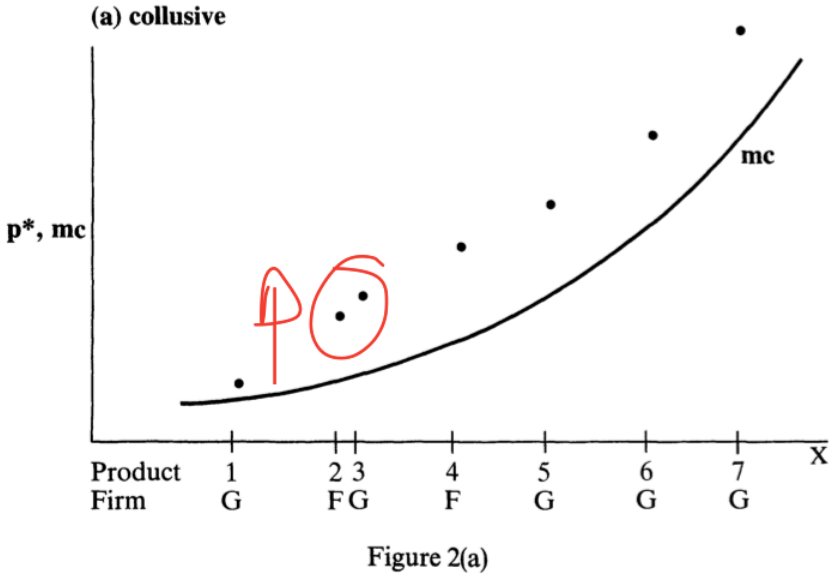


Figure 2(a)

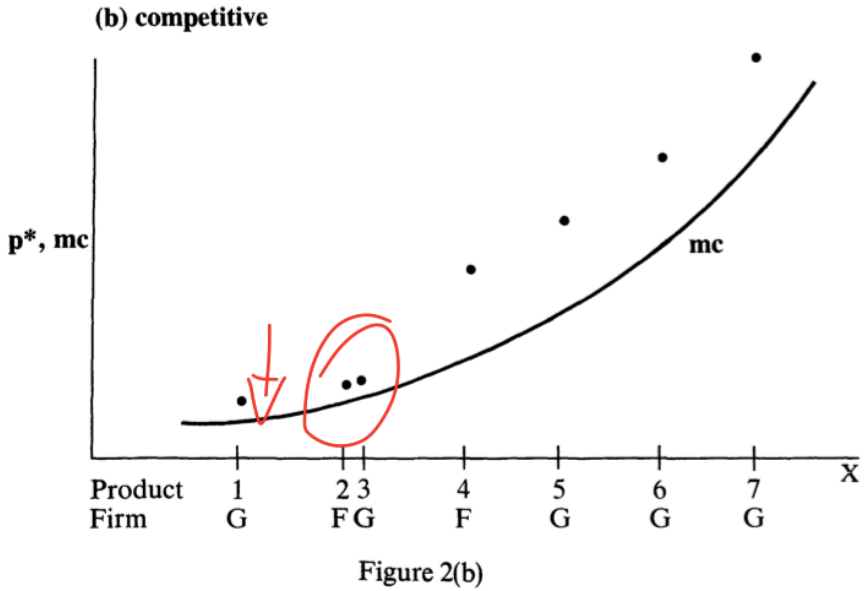


Figure 2(b)

- If $\theta^0 = \text{coll.}$, $\pi_{jt}^{\text{comp.}} - \pi_{jt}^0$ is correlated with the degree of product differentiation **across** products owned by G and F (i.e. residual demand **rotation** for Ford).
- **Proposed IV:** Distribution of characteristic differences between products owned by the same firm, and products owned by competitors products.
- **Weak IV:** Aggregate market attributes will fail to explain margin differences across products, and will not be able to identify the correct conduct model.

Alternative approach: Direct estimation

Gandhi and Houde (ongoing)

- The supply relation is an *unknown* function of the *known* distribution of substitution patterns (Δ) and the *unknown* "coalitions" of products (Θ):

$$p_{jt} = \text{MC}_{jt} + \pi_j(\Delta_t, \Theta_t)$$

- In practice the equilibrium markup function does not have to satisfy a specific FOC. For instance,
  - In Canada, gasoline stations usually set prices using 1 cents/liter intervals (i.e. $1c/l \approx 20\%$ of profit margins).
  - Clark and Houde (2012): Stations were colluding on a **uniform** price (i.e. between 20 and 80 stations)
  - Miller and Weinberg (2015): ABI and Miller/Coors increase prices by **identical** magnitudes post-merger.
- An alternative to the FOC approach is to *approximate* the supply relation using estimates of the elasticity of substitutions and known/suspected coalitions.

An Illustration: New car market

- With the demand estimates, we can compute the degree of substitution between products using **diversion-ratios**:

$$\text{DR}_{j,k} = \frac{\partial D_k / \partial p_j}{-\partial D_j / \partial p_j}$$

- The relative isolation of product j is,

$$C_j(r) = \sum_{k=1}^{r} \text{DR}_{j,(k)}$$

where $\text{DR}_{j,(k)}$ is the k^{th} largest diversion-ratio.

An Illustration: New car market

- With the demand estimates, we can compute the degree of substitution between products using **diversion-ratios**:

$$\text{DR}_{j,k} = \frac{\partial D_k / \partial p_j}{-\partial D_j / \partial p_j}$$

- The relative isolation of product j is,

$$C_j(r) = \sum_{k=1}^{r} \text{DR}_{j,(k)}$$

where $\text{DR}_{j,(k)}$ is the k^{th} largest diversion-ratio.

- **Competition vs Collusion**:

$$\ln p_{jt} = W_{jt}\gamma + \lambda_1 \ln C_{jt}^{F}(r) + \lambda_2 \ln C_{jt}^{-F}(r) + \omega_{jt}$$

where $C_{jt}^{-F}(r)$ and $C_{jt}^{F}(r)$ measures the relative isolation of j relative to products in and out of network F.

Are car makers internalizing substitution patterns?

	(1) OLS	(2) OLS	(3) IV	(4) IV
$C^{-F}(1)$	-0.18[a]		-0.09[a]	
	(0.01)		(0.02)	
$C^{F}(1)$	0.0004		0.008	
	(0.003)		(0.00863)	
$C^{-F}(10)$		-0.22[a]		-0.11[a]
		(0.02)		(0.03)
$C^{F}(10)$		0.001		0.005
		(0.003)		(0.009)
R^2	0.926	0.928	0.922	0.923

Instrumental variables:

- $\mathbf{C}^{F}(r)$: Histogram of characteristics differences, and sum of local competitor characteristics (i.e. Z_{jt})
- $\mathbf{C}^{-F}(r)$: Average distance between products owned by the same firm

Reduced-Form Approach

- The structural approach has been criticized for its reliance on parametric assumptions (e.g. mixed-logit) and strict equilibrium assumptions (e.g. Bertrand-Nash).
  - See Angrist and Pischke (2010) for a discussion.
- An alternative approach is to rely on *plausibly exogenous* changes in conduct, to test wether or not firms exert market-power.
  - **Difference-in-Difference:** What is the effect of a conduct change on the price of *treated* markets, relative to prices in *control* markets.
- Examples of quasi-experiments related to industry conduct:
  - *Vertical merger in gasoline markets:* Hastings (2004), Houde (2012)
  - *Horizontal mergers:* Ashenfelter, Hosken, and Weinberg (2015), Miller and Weinberg (2016), Allen, Clark and Houde (2014)
  - *Price fixing inquiry:* Clark and Houde (2014)
  - Many others...

A Primer on Treatment Effect Estimation

- In economics, we are often interested in estimating the average effect of a discrete policy change (D_i):

$$\alpha = E(Y_i|D_i = 1) - E(Y_i|D_i = 0)$$

- Main challenge: Treatment is not randomly assigned to individuals

$$E(Y_i|D_i = 1) - E(Y_i|D_i = 0) = \underbrace{E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1)}_{\text{Avg. treatment effect on treated}} + \underbrace{E(Y_{0i}|D_i = 1) - E(Y_{0i}|D_i = 0)}_{\text{Selection bias}}$$

  where Y_{ji} is the outcome of individual i with treatment $j = 0, 1$.

- Useful references:
  - Fabien Waldinger's notes: www.fabianwaldinger.com/applied-econometrics
  - Chris Taber's notes: www.ssc.wisc.edu/∼ctaber/706/diffdiff.pdf

A Primer on Treatment Effect Estimation (continued)

- To fix ideas, consider the following regression model:

$$Y_i = x_i b + \alpha D_i + e_i$$

- In general we have three ways of dealing with this problem:
  - Conduct field/lab experiments that satisfy the random assignment assumption
  - Instrumental variables: $E(e_i|x_i, z_i, D_i = 1) = E(e_i|x_i, z_i, D_i = 0) = 0$ and $corr(z_i, D_i)$ is large.
  - "Fixed-effect" methods: $E(e_i|x_i, \mu_i, D_i = 1) = E(e_i|x_i, \mu_i, D_i = 0) = 0$
- The later approach consists of assuming that the treatment is randomly assigned conditional on individuals/markets with similar characteristics.
- Examples of fixed-effect approaches:
  - Propensity score matching (or weighted regression)
  - Difference-in-difference
  - Regression discontinuity

A Primer on Treatment Effect Estimation (continued)

- **Diff-in-Diff:** Standard linear fixed-effect regression
  - Three-way panel: Individuals (i), groups (s), and time (t)
  - Treatment indicator: $D_{ist} = 1(s \in \text{Treatment group}) \times 1(t > T_s)$
- **Regression line:**

$$Y_{ist} = \mu_s + \gamma_t + \alpha D_{ist} + x_{ist} b + e_{ist}$$

Let G denotes the number of groups (s), and T the number periods.

- **Assumption:** Parallel trends and time-invariant group differences

$$E(Y_{ist}|s,t) = \mu_s + \gamma_t$$

and

$$E(e_{ist}|x_{ist}, \mu_s, \gamma_t, D_{ist} = 1) = E(e_{ist}|x_{ist}, \mu_s, \gamma_t, D_{ist} = 0) = 0$$

Limitations of Diff-in-Diff

- The main identifying assumptions (pre-trend) is difficult to test
  - The best thing to do is to report differences in pre-treatment trends across the two groups (imperfect).
- Estimation of the treatment effect often relies on a **small** number of policy change
  - Identification of α is at the **group** level (s) not at the individual level.
  - Standard-errors must be clustered at the group level to reflect this fact
  - Often G is very small: $\downarrow$ the precision of the estimates, and $\uparrow$ sensitivity to outliers.
- The "common-trend" assumption has important economic implications:
  - The parameter γ_t is used to predict what would have happened to individuals in the treatment group, absent of the policy change.
  - A "good" control group is a set of individuals affected in the same way by **aggregated shocks**, but are not directly affected by the treatment.
  - **Challenge in IO**: Firms in the same market can be affected by D_t, and firms outside of the market might not be good controls $\gamma_t \neq \gamma'_t$.

Mergers as Quasi-Experiments

Hastings, AER, 2004

- **Question:** What is the impact of vertical integration (VI) on gasoline prices?
- "Model-free" answers:
  - *Cross-sectional comparison:* Price differences between VI and non-VI stations.
  - *Time-series comparison:* Price changes before/after a vertical merger.
- **Simultaneity problems:** Both estimates are biased due to the endogeneity of vertical contracts
  - VI contracts are not randomly assigned
  - The timing of mergers is correlated with aggregate shocks
- **Solution:** We need to find a group of *control* products or markets.
- **Control group requirements:**
  - *Plausibly* unaffected by the observed change in conduct (e.g. merger),
  - Share the common market features, in order to control for *aggregate shocks* correlated with the change in conduct.

Difference-in-Difference Regression

- **Differentiation assumption:** Stations located more than $\bar{d}$ miles away from each other face zero cross-elasticity of substitution.
- **Implication:** The price effect of a vertical merger is *localized* around the location of stations experiencing a change in vertical contract.
  - *Case study:* Acquisition of Trifty stations (independent) by ARCO (vertically integrated).
- **Notation:**
  - p_{jt} is the price posted at station j in period t
  - $D_j = 1$ is station j is located within $\bar{d}$ miles of a "treated" station
  - $T_t = 1$ in post-merger period
- Diff-in-Diff regression:

$$p_{jt} = \mu_j + \tau_t + \lambda T_t \times D_j + u_{jt}$$

where λ measures the average treatment effect of the merger on stations located near a Trifty station.

Estimation Results

- **Key result:** Competing with an independently operated station *reduce* prices by 5 cents/gallon.
- *Coordinated effects:* With resale-price maintenance, integrated networks internalize cannibalization effects.
- *Quality effects:* The rebranding of independent stations increases WTP.

Dependent variable: Retail price for regular unleaded			
Variable	(1)	(2)	(3)
Intercept	1.3465	1.3465	1.3617
	(0.0421)	(0.0415)	(0.0287)
Company operated	0.1080	−0.0033	−0.0033
	(0.0107)	(0.0178)	(0.0122)
Independent	—	−0.1013	−0.0500
		(0.0143)	(0.0101)
LA∗February	—	—	0.0180
			(0.0065)
LA∗June	—	—	0.0243
			(0.0065)
LA∗October	—	—	0.1390
			(0.0064)
SD∗February	—	—	−0.0851
			(0.0036)
SD∗June	—	—	−0.0304
			(0.0036)
SD∗October	—	—	0.0545
			(0.0036)
Adjusted R^2	0.3772	0.3953	0.7181
F-test for no fixed effects:			
Numerator DF: 668			
Denominator DF: 1,999			
F-value: 3.262			Prob. > F: 0.000
Hausman test for random effects:			Prob. > M: 0.000
Hausman's M value: 622.296			

Merger Retrospective as a Validation Tool

Houde, AER, 2012

- The Hotelling model
  - Consumers have an "ideal" point (or address) to purchase gasoline, and tradeoff the transportation cost and potential price discount of traveling further.
  - *Implication:* The price effect of mergers is zero for stations located far from the merging firms (i.e. control group).
- Incorporating driving into the model:
  - *Multiple ideal points:* Everything else being equal, consumers are indifferent between stations located along their driving path.
  - *Transportation cost*: Deviation from commuting path
  - *Implication:* Stations compete with their immediate neighbors, AND with stations located along common paths.
- Empirical application: Québec City market between 1991-2001
  - Lots of commuters: Limited public transit and concentrated job locations
  - Vertical merger: Acquisition by Ultramar of Sunnoco stations in 1997.

Demand Estimation

- **Demand estimation:**
  - Structurally estimate the price sensitivity and transportation cost tradeoff using Berry et al. (1995) framework
  - *Exogenous variation:* Progressive exit of small-scale stations due to technology change ($\approx 30\%$)
  - Controlling for location fixed-effects, changes in the number of local competitors (non-linearly) identifies the travel cost.
- **Key results:**
  - Ignoring commuting patterns severely biased travel cost estimates
  - Why? High demand stations are located at the intersection of major commuting paths, far from where people live (home address)
  - Despite large travel costs, the model with commuting implies low market-power due to the connectivity of the road network
  - Average profit margins are estimated to be 8%
- Implications for merger analysis:
  - Standard control group definition produces a downward bias estimate of the merger effect. Why?
  - Distant competitors sharing common commuters are substitute.

Merger effect: Reduced-form *vs* Model simulation

- Retrospective analysis:
  - Diff-in-Diff regression:

$$\begin{aligned} p_{jt} \quad &= \mu_j + \tau_t + \lambda_1 D_j \times T_t \times 1(j \in \text{Ultramar}) \\ &\quad + \lambda_2 D_j \times T_t \times 1(j \notin \text{Ultramar}) + u_{jt} \end{aligned}$$

  - λ_1 measures the net effect of the merger on stations operated by Ultramar (vertical chain)
  - λ_2 measures the net effect of the merger on competing chains
- Merger simulation:
  - **Assumption:** Ultramar sets prices at all of its stations (i.e. resale-price maintenance).
  - With the demand estimate, we can infer the marginal-cost of each stations **before** the merger:

$$\text{mc}_{jt} = p_{jt} - \Sigma_{jt}^{-1} s_{jt}$$

    where $\Sigma_{jt} = \Delta_{jt} \cdot \Theta_t^{before}$.
  - We can then solve the **post-merger** equilibrium prices by using a new conduct matrix: Θ_t^{after}.

Validation Results

- **Key results:**
  - The "direct" effect of the merger on Ultramar stations is large and very localized.
  - The "indirect" effect is very small
  - The reduced-form approach leads to small under-estimate of the direct effect

		Merger simulation	Retrospective estimates
Panel A: Aggregate price effects			
Merger effects	Treated	0.406	–
	Control	0.019	–
	Difference	*0.388*	*0.433*
Panel B: Direct/Indirect decomposition			
Sunoco price change		0.607	–
Ultramar network	Treated	0.523	–
	Control	0.101	–
	Difference	*0.422*	*0.501*
Competing networks	Treated	0.026	–
	Control	0.000	–
	Difference	*0.025*	*0.0384*

Notes: Treatment group includes stores located within 1/2 minutes of a Sunoco station. Prices are expressed in cents per liter.

Market Power in Price Negotiation Markets

Allen, Clark and Houde, AER, 2014

- Retrospective merger analysis is often useful when we know very little about the possible conduct models.
- **Example:** What is the effect of competition in markets with search frictions?
  - Theoretical search models have ambiguous predictions
  - Standard models are not directly applicable to settings in which search frictions arise due to *negotiation costs*
- **Context:** The Canadian mortgage market
  - Homogenous contracts fully insured by the Government
  - Prices are negotiated with local branch officers
  - *Price dispersion:* (i) 25% pay a common posted-rate, (ii) average discount is 75% basis-points, and (iii) residual dispersion accounts for about 60% of the variance of negotiated rate spreads.
- **Conduct change:** Merger between one of the largest bank, and a large trust company.

Reduced-form effect of the merger

- **Assumption:** The choice-set of consumers is defined by the number of lenders located with 5 KM of the new home.
- **Implications:**
  - *Treatment group:* Neighborhoods with both lenders before the merger experienced a reduction in competition.
  - *Control group:* Neighborhoods with one or none of the lenders faced the same level of competition before/after.
- This lead to a natural diff-in-diff regression:

$$r_i = \mu_{\text{nbh}(i)} + \tau_{\text{week}(i)} + \lambda D_i \times T_i + u_i$$

  where i indexes a contract, μ and τ are neighborhood and week fixed-effects, and $D_i \times T_i = 1$ in treated neighborhoods after the merger.
- Two issues:
  - Treated neighborhoods are more competitive than control groups (at least two)
  - If price discrimination is present, the effect should be *heterogeneous*

Average Treatment Effects

- **Sample:** Neighborhoods with intermediate levels of competition (i.e. 5-8 lenders)
- **Results:** Eliminating one potential lender,
  - Increases the probability of receiving no discounts by 5%
  - Reduces the average discount by 7 bps.

	Margin		Zero discount	
	Baseline	With trend	Baseline	With trend
Linear DiD (OLS)				
Merger ATE	0.0607^a	0.0719^a	0.0646^a	0.0477^a
	(0.0183)	(0.0242)	(0.0154)	(0.0188)
N	18,121	18,121	18,121	18,121
R^2	0.408	0.420	0.181	0.189
Matching DiD				
Merger ATE	0.0739^a	0.0692^b	0.0670^a	0.0538^b
	(0.0249)	(0.0273)	(0.0206)	(0.0218)
N	17,220	17,220	17,220	17,220
Change-in-change				
Merger ATE	0.057^a	0.0661^a		
	(0.015)	(0.021)		
N	18,103	18,103		

What is the effect of the merger on the distribution of negotiated prices?

- **Athey and Imbens (2002):** Non-linear diff-in-diff
  - *Goal:* Measure the effect of the merger along every percentiles of the residual rate distribution.
- How does it work?
  - The *counter-factual* price distribution in treated markets can be recovered from three *known* distributions:

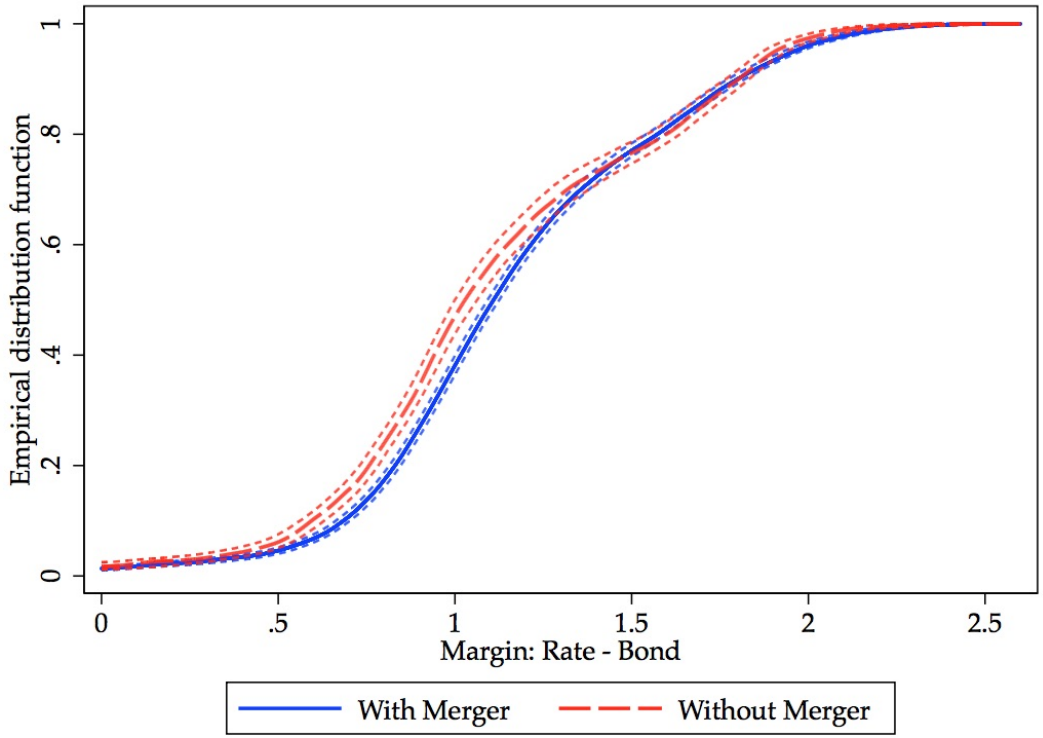
$$F^c_{1,1}(p) = F_{1,0}\left(F^{-1}_{0,0}\left(F_{0,1}(p)\right)\right)$$

    where $F_{g,t}(p)$ is the CDF in group $g \in (0,1)$ in period $t \in (0,1)$.
  - Intuitively, this equation predicts the quantile change that would have been observed in *treated* markets by observing the quantile change in the control markets.
  - This allows us to measure the merger effect at every percentiles u_i:

$$\alpha(u_i) = F^{-1}_{1,1}(u_i) - F^c_{1,1}(u_i)$$

What is the effect of the merger on the distribution of negotiated prices?

- **Good negotiators:** Strong effects at the bottom.
- **Poor negotiators:** Zero effects at the top.
- An **increase** in competition leads to an **increase** in price dispersion
  - *Competitive* price discrimination.



Conclusion

- Reduced-form *vs* structural approach:
  - RF can be very useful to test specific conduct assumptions, without relying on functional forms or model assumptions.
  - However, it does not fully answer our original question
    - ★ Aggregate effects of merger are not identified (only difference)
    - ★ Some conduct models are indistinguishable: Mergers in competitive **and** collusive industries have zero effects.
- What's next?
  - We still have incomplete tools to identify firm conduct in differentiated product markets
  - Some market structures are understudied: search markets, business-to-business markets, two-sided markets, asymmetric information, etc.
  - Need more applications!

Byrne, D. and N. de Roos (2017, October).
Learning to coordinate: A study in retail gasoline.
Working paper, University of Melbourne.

Conley, T. and C. Taber (2011).
Inference with "difference in differences" with a small number of policy changes.
Review of Economics and Statistics.