Diversion Ratios

Chris Conlon

September 24, 2020

Grad IO

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 (2013/2015/2020...))
- Possibly more empirical experience at DOJ/FTC.

Where do Diversion Ratios come from?
(Stolen from Conlon and Mortimer (2020)

In Theory

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

$$\arg\max_{p_j} (p_j - c_j) \cdot q_j(p_j, p_{-j}) + (p_k - c_k) \cdot q_k(p_j, p_{-j})$$

$$0 = q_j + (p_j - c_j) \cdot \frac{\partial q_j}{\partial p_j} + (p_k - c_k) \cdot \frac{\partial q_k}{\partial p_j}$$

$$p_j = -q_j / \frac{\partial q_j}{\partial p_j} + c_j + (p_k - c_k) \cdot \underbrace{\frac{\partial q_k}{\partial p_j} / - \frac{\partial q_j}{\partial p_j}}_{D_{jk}}$$

$$p_j = \underbrace{\frac{\epsilon_{jj}}{\epsilon_{jj} + 1}}_{\text{Lerner Markup}} \left[c_j - \overbrace{\underbrace{c_j \cdot e_j}_{\text{efficiency}} + \underbrace{(p_k - c_k) \cdot D_{jk}(p_j, p_{-j})}_{\text{opp cost}}}^{UPP} \right]$$

Caveat: UPP, Partial Merger, Full Merger.

In Theory

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

$$\arg\max_{p_j} (p_j - c_j) \cdot q_j(p_j, p_{-j}) + (p_k - c_k) \cdot q_k(p_j, p_{-j})$$

$$0 = q_j + (p_j - c_j) \cdot \frac{\partial q_j}{\partial p_j} + (p_k - c_k) \cdot \frac{\partial q_k}{\partial p_j}$$

$$p_j = -q_j / \frac{\partial q_j}{\partial p_j} + c_j + (p_k - c_k) \cdot \underbrace{\frac{\partial q_k}{\partial p_j} / - \frac{\partial q_j}{\partial p_j}}_{D_{jk}}$$

$$p_j = \underbrace{\frac{\epsilon_{jj}}{\epsilon_{jj} + 1}}_{\text{Lerner Markup}} \left[c_j - \overbrace{\underbrace{c_j \cdot e_j}_{\text{efficiency}} + \underbrace{(p_k - c_k) \cdot D_{jk}(p_j, p_{-j})}_{\text{opp cost}}}^{UPP} \right]$$

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:

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

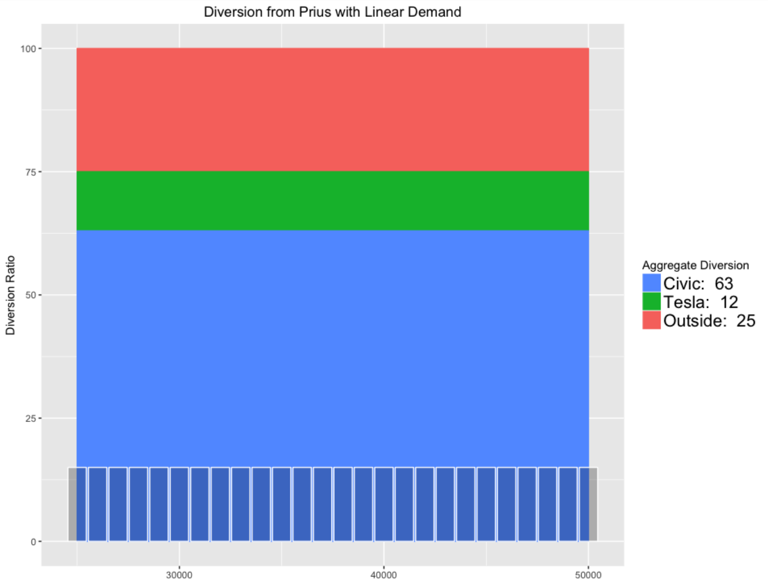
- Diversion, $D_{jk}(p_j, p_{-j})$, is a Marginal Treatment Effect (MTE) in the language of Heckman and Vytlacil (1999).
- It is a function. Actually a matrix valued function.
- It is not identified non-parametrically from a single price increase.

Various Treatment Effects

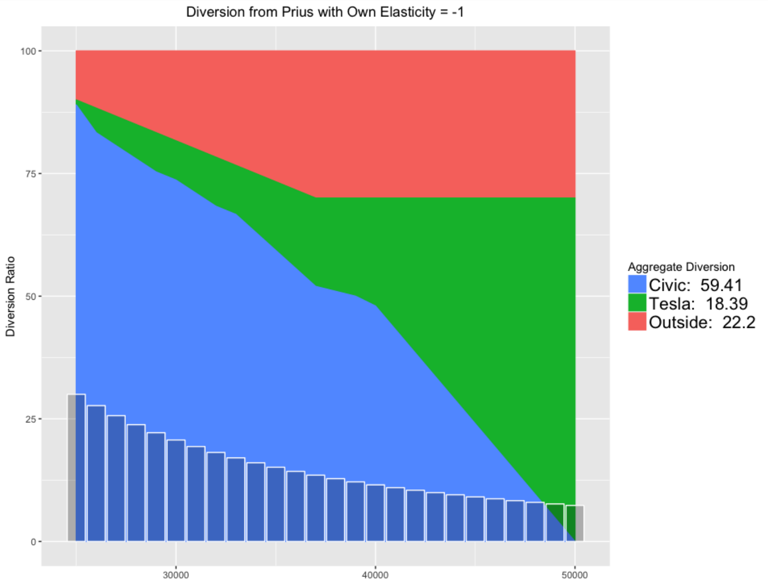
- Determine what different measures of diversion identify.
  - Finite price increase $\rightarrow$ local average treatment effect (LATE)
  - Product removal (treating everyone) $\rightarrow$ average treatment effect (ATE)
  - A nonparametric function of p_j $\rightarrow$ marginal treatment effect (MTE)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

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

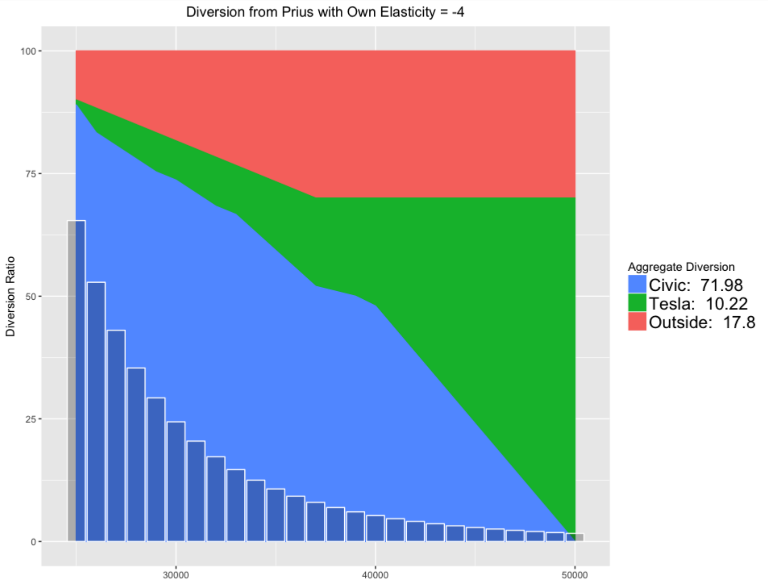
Thought Experiment – Linear Demand for a Toyota Prius



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

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

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)

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

  - 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?