Diversion Ratios

Chris Conlon

Wednesday 15th October, 2025

Grad IO

What are Diversion Ratios?

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?

$$D_{j \to k}(p_j, \mathbf{p}_{-j}) = \frac{\frac{\partial q_k}{\partial p_j}}{\left| \frac{\partial q_j}{\partial p_j} \right|}$$

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 (2021))
- Commonplace at DOJ/FTC.

Nevo (2000) Example

#	Brand	K Corn Flakes	K Raisin Bran	K Frosted Flakes	K Rice Krispies
1	K Corn Flakes	−3.696	.023	.500	.010
2	K Raisin Bran	.023	−2.061	.088	.051
3	K Frosted Flakes	.361	.059	−3.546	.028
4	K Rice Krispies	.010	.048	.040	−1.320
5	K Frosted Mini Wheats	.000	.053	.003	.057
6	K Froot Loops	.000	.010	.008	.038
7	K Special K	.155	.072	.248	.039
8	K NutriGrain	.270	.094	.313	.023
9	K Crispix	.003	.038	.020	.079
10	K Cracklin Oat Bran	.000	.023	.001	.046
11	GM Cheerios	.007	.080	.035	.069
12	GM Honey Nut Cheerios	.001	.017	.017	.043
13	GM Wheaties	.503	.113	.445	.029
14	GM Total	.140	.064	.238	.042
15	GM Lucky Charms	.000	.012	.010	.041
16	GM Trix	.000	.010	.009	.043
17	GM Raisin Nut	.007	.137	.043	.059
18	P Raisin Bran	.014	.232	.063	.050
19	P Grape Nuts	.001	.048	.006	.050
20	Q 100% Natural	.000	.023	.002	.048
21	Q Life	.003	.038	.052	.048
22	Q CapNCrunch	.001	.013	.015	.038
23	R Chex	.005	.037	.028	.081
24	N Shredded Wheat	.002	.081	.018	.049

Advantages of Diversion

- When you were taught elasticities you were taught they were a unit free comparison across products, markets, etc.
- But that is only true of own elasticities not cross elasticities.
- Is $\epsilon_{jt} = .01$ or $\epsilon_{jt} = .03$ a better substitute? We can't tell.
  - You need to take $\epsilon_{jk} \cdot s_k$ to know. If you take $\epsilon_{jk} \cdot \frac{s_k}{p_j} = D_{jk}$ (back at diversion).
  - Because it tells us fraction of switchers choosing k we can compare across settings (since share sums to one).
  - ie: diversion ratio $D_{jk} = 0.1$ actually tells me something!
- Whenever you are tempted to report cross elasticities report diversion ratio instead.

	Cheerios	Special K	Corn Flakes	Reese’s Puffs	Capt Crunch	Froot Loops	Shares
HN Cheerios	5.07	4.27	3.75	5.33	3.58	3.48	2.69
Frosted Flakes	2.46	2.54	4.54	4.00	5.35	7.24	2.65
Cheerios	-2.46	5.91	3.13	3.19	1.36	1.77	2.10
Honey Bunches	2.47	2.51	2.21	2.08	1.94	1.99	1.47
Cinn Toast Crunch	3.43	2.10	1.69	3.00	1.78	1.84	1.43
Froot Loops	1.26	1.19	1.64	1.69	1.82	-2.71	1.18
Lucky Charms	2.18	1.64	1.57	2.99	1.59	1.58	1.14
Frosted Mini-Wheats	0.36	0.50	0.74	0.68	0.87	1.27	1.01
Corn Flakes	2.01	2.18	-2.64	1.31	1.24	1.52	0.98
Rice Krispies	1.50	1.72	1.56	0.89	0.68	1.25	0.96
Apple Jacks	0.91	0.80	1.24	1.27	1.42	2.45	0.85
Raisin Bran (KEL)	0.46	0.47	0.63	0.78	0.82	1.24	0.79
Special K Red Berry	0.96	1.45	0.95	0.78	0.68	0.90	0.75
Special K	2.06	-2.66	1.18	0.71	0.44	0.58	0.74
MG Cheerios	1.11	0.99	0.75	0.89	0.54	0.66	0.71
Reese’s Puffs	1.36	0.86	0.87	-2.70	1.08	1.01	0.69
Life	1.15	1.12	1.05	1.02	1.72	0.89	0.68
Cocoa Puffs	1.18	0.92	0.95	1.47	1.05	0.97	0.67
Capt Crunch	0.63	0.58	0.88	1.21	-2.68	1.19	0.62
Capt Crunch Berry	0.68	0.61	0.83	1.15	3.29	1.00	0.58
Corn Pops	0.43	0.43	0.71	0.66	0.75	1.45	0.56
Cinn Life	0.76	0.75	0.83	0.84	1.59	0.78	0.54
Fruity Pebbles	0.61	0.59	0.71	0.71	0.75	0.77	0.44

Where do Diversion Ratios come from?
(Stolen from Conlon and Mortimer (2021)

In Theory

Consider Bertrand FOC's for multi-product firm j:

$$\rightarrow p_j = q_j(\mathbf{p}) \left[-\frac{\partial q_j}{\partial P_j}(\mathbf{p}) \right]^{-1} + c_j + \sum_{k \in \mathcal{J}_f \setminus j} (p_k - c_k) \underbrace{\frac{\partial q_k}{\partial P_j}(\mathbf{p}) \left[-\frac{\partial q_j}{\partial P_j}(\mathbf{p}) \right]^{-1}}_{D_{j \to k}(\mathbf{p})}$$

$$p_j(p_{-j}) = \underbrace{\frac{1}{1 + 1/\epsilon_{jj}(\mathbf{p})}}_{\text{Markup}} \left[c_j + \sum_{k \in \mathcal{J}_f \setminus j} (p_k - c_k) \cdot D_{j \to k}(\mathbf{p}) \right].$$

Mult-product pricing raises the opportunity cost of selling j.

Agencies often calculate Upward Pricing Pressure or UPP asks how merger changes as we combine $\mathscr{J}_f$ and $\mathscr{J}_g$:

$$\left[c_j + \sum_{k \in \mathscr{J}_f \setminus j} (p_k - c_k) \cdot D_{j \to k}(\mathbf{p})\right]$$
$$UPP_j = \Delta c_j + \sum_{k \in \mathscr{J}_g} (p_k - c_k) \cdot D_{j \to k}(\mathbf{p})$$

How does the merger change the opportunity cost for j?

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

$$UPP_j \approx (p - mc) \sum_k D_{j\to k}(\mathbf{p}) - \Delta mc_j$$

If several brands of acquisition have the same markup – can consider firm-level diversion. (We can aggregate diversion across similar flavors)

Ignoring Efficiencies

$$GUPPI_j \approx \frac{(p_j - mc_j)}{p_j} D_{j\to k}(\mathbf{p})$$

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.

Wald Estimator

The diversion ratio $D_{j \to k}(p_j, x) \equiv \frac{\partial q_k}{\partial P_j}(p_j, x) / - \frac{\partial q_j}{\partial P_j}(p_j, x)$ can be obtained as the limit of the Wald estimator where the price increase (or decrease) becomes small, so long as demand slopes *strictly* downwards $\frac{\partial q_j}{\partial P_j} < 0$.

$$\lim_{p'_j \to p_j} \frac{q_k(p'_j, x) - q_k(p_j, x)}{-(q_j(p'_j, x) - q_j(p_j, x))} \to \frac{\frac{\partial q_k}{\partial P_j}(p_j, x)}{-\frac{\partial q_j}{\partial P_j}(p_j, x)} \equiv D_{j \to k}(p_j, x)$$

Analogue to LATE Theorem Imbens Angrist (1994)

Theorem (Conlon Mortimer (2021))

Under the following conditions:

(a) Mutually Exclusive and Exhaustive Discrete Choice: $d_{ij} \in \{0,1\}$ *and* $\sum_{j \in \mathcal{J}} d_{ij} = 1$.

(b) Exclusion: $u_{ik}(p_j, x) = u_{ik}(p'_j, x)$ *for all* $k \neq j$ *and any* (p_j, p'_j)*;*

(c) Monotonicity: $u_{ij}(p'_j, x) \leq u_{ij}(p_j, x)$ *for all* i *and any* $(p'_j > p_j)$*; and*

(d) Existence of a first-stage: $Pr(d_{ij}(p_j, x) = 0) \neq Pr(d_{ij}(p'_j, x) = 0)$ *for* $(p'_j > p_j)$*;*

(e) Random Assignment: $(u_{ij}(P_j, x), u_{ik}(P_j, x)) \perp P_j$.

then the Wald estimator:

$$\frac{q_k(p'_j, x) - q_k(p_j, x)}{-\left(q_j(p'_j, x) - q_j(p_j, x)\right)} = \mathbb{E}[D_{jk,i}(x) | d_{ij}(p_j, x) > d_{ij}(p'_j, x)]$$

Diversion as Treatment Effects

Outcome $Y_i \in \{0,1\}$ denotes the event that consumer i purchases product k: $d_{ik}(P_j) = 1$.

Treatment $T_i \in \{0,1\}$ denotes the event that consumer i does **not** purchase product j. In other words $T_i = 0$ implies $d_{ij}(P_j) = 1$ and $T_i = 1$ implies $d_{ij}(P_j) = 0$.

Instrument $Z_i = P_j$ the price of j induces consumers into not purchasing j.

Diversion as MTE

What can we learn from a change in p_j?

$$\text{Wald}(p_j, p_j', x) = \int_{p_j}^{p_j'} D_{jk}(p_s, x) w(p_s)\, \partial\, p_s \text{ with } w(p_s) = \frac{\frac{\partial q_j(p_s, x)}{\partial p_j}}{\int_{p_j}^{p_j'} \frac{\partial q_j(p_t, x)}{\partial p_j} \partial p_t}$$

$$= \int_{p_j}^{p_j'} \int D_{jk,i}(p_s, x) w_i(p_s, x)\, \partial p_s\, \partial F_i \quad \text{with } w_i(p_s, x) = \frac{\left|\frac{\partial q_{ij}(p_s, x)}{\partial p_j}\right|}{q_j(p_j, x) - q_j(p_j', x)}$$

$$= \int D_{jk,i}(x)\, w_i(z_j, z_j', x)\, \partial F_i \quad \text{with } w_i(z_j, z_j', x) = \frac{q_{ij}(z_j, x) - q_{ij}(z_j', x)}{q_j(z_j, x) - q_j(z_j', x)}$$

- Individual diversion ratios in logit family are $D_{jk,i} = \frac{s_{ik}}{1 - s_{ij}}$ and don't depend on p_j, z_j
- Which individuals respond to z_j or p_j determines the weighting scheme only.

Weights

Again $D_{jk,i} = \frac{s_{ik}}{1-s_{ij}}$ for logit family:

	$w_{ij}(x) \propto$
second choice data	$s_{ij}(x)$
price change $\frac{\partial}{\partial p_j}$	$s_{ij}(x) \cdot \big(1 - s_{ij}(x)\big) \cdot \lvert\alpha_i\rvert$
characteristic change $\frac{\partial}{\partial x_j}$	$s_{ij}(x) \cdot \big(1 - s_{ij}(x)\big) \cdot \lvert\beta_i\rvert$
small quality change $\frac{\partial}{\partial \xi_j}$	$s_{ij}(x) \cdot \big(1 - s_{ij}(x)\big)$
finite change in z_j	$s_{ij}(z_j') - s_{ij}(z_j)$

- Normalize so that above weights sum to one $w_{ij}^* = w_{ij} / \int w_{ij} \, dF_i$
- For plain logit $D_{jk,i} = \frac{s_k}{1-s_j}$ for all i and weights don't matter!

Here we get a really simple decomposition:

$$D_{j\to k} = \sum_{i\in\mathcal{J}} \underbrace{\frac{s_{ik}}{1-s_{ij}}}_{D_{jk,i}} \cdot \underbrace{\frac{\pi_i \cdot s_{ij}}{s_j}}_{\gamma_{ij}}$$

Idea: different interventions change weights but not individual diversion ratios:

- **Individual Diversion Ratios** $D_{j\to k,i}$: Pr$\left(i \text{ chooses } k \mid \text{ doesn't choose } j\right)$.
- **Weights** γ_{ij} : For second choices: Pr$\left(\text{Type } i \mid \text{ leaves } j\right) = $ Pr$\left(\text{Type } i \mid \text{ chooses } j\right)$.

  For infinitesimal price changes $\gamma_{ij} = \frac{\pi_i \cdot |\beta_i^p| \cdot s_{ij}(1-s_{ij})}{\sum_{i\in\mathcal{J}} \pi_{i'} \cdot |\beta_i^p| \cdot s_{i'j} \cdot (1-s_{i'j})}$ or finite: $\gamma_{ij}(p_j, p_j') = \frac{s_{ij}(p_j')-s_{ij}(p_j)}{s_j(p_j')-s_j(p_j)}$.

  (Types that are: price sensitive and likely to choose j).

  But we could do this with anything (not just price!)

Application of CM Decomposition: Einav, Guido, Klenow (2025)

Measuring customer overlap by counting total sales of substitute k (Dunkin) and purchases of NOT focal product j (Starbucks) using panel data:

$$C_{j\to k} = \frac{\sum_{\{it|i\in\mathcal{C}_j\}} 1\left(d_{it}=k\right)}{\sum_{\{it|i\in\mathcal{C}_j\}} 1\left(d_{it}\neq j\right)} = \frac{\sum_{i\in\mathcal{C}_j} q_{i,k}}{\sum_{i\in\mathcal{C}_j} q_{i,(-j)}} = \frac{Q_k}{\left(\sum_{i\in\mathcal{C}_j} T_i\right)-Q_j} \approx \frac{s_k}{1-s_j} \;\Big|\; \text{ever chooses } j$$

- Restrict to the set $\mathcal{C}_j$ of households that ever purchased j.
- T_i are total purchases in category by household i.
- For any full support random utility model $u_{ij}=v_{ij}(\theta_i)+\varepsilon_{ijt}$ as $T_i\to\infty$.
  i will always choose j with positive probability. This will make the $\mathcal{C}_j$ restriction meaningless.
- In the limit we should get the logit prediction.

Application of CM Decomposition: Einav, Guido, Klenow (2025)

Can I rewrite this as a weighted average of individual diversion ratios? (Yes)

$$C_{j\to k} = \frac{\sum_{i\in\mathcal{C}_j} q_{i,k}}{\sum_{i\in\mathcal{C}_j} q_{i,(-j)}} = \frac{\sum_{i\in\mathcal{C}_j} T_i \cdot \hat{s}_{ik}}{\sum_{i\in\mathcal{C}_j} T_i \cdot (1-\hat{s}_{ij})} = \sum_{i\in\mathcal{C}_i} \underbrace{\frac{\hat{s}_{ik}}{1-\hat{s}_{ij}}}_{\hat{D}_{j\to k,i}} \cdot \underbrace{\frac{q_{i,(-j)}}{\sum_{i'\in\mathcal{C}_j} q_{i',(-j)}}}_{=\gamma_{ij}}.$$

- Define: $\hat{s}_{ij} = \frac{1}{T_i}\sum_t 1(c_{it}=j)$, and $q_{i,(-j)} = \sum_t 1[c_{it}\neq j] = T_i(1-\hat{s}_{ij})$.
- Now the weights are Pr(Type i | chooses NOT j) for i where $\hat{s}_{ij} > 0$.
- We put the most weight on people who have the most non Starbucks visits (but zero if they never visit).
- But with individual data, we could re-weight the households in the sample γ_i such that we get second-choice diversion (or whatever else we want).

Welfare and WTP

As we learned on HW2, welfare and willingness to pay are mostly about diversion to outside good

$$WTP_i(j) = \mathbb{E}[\max_{k \in \mathcal{J}} u_{ik}] - \mathbb{E}[\max_{k' \in \mathcal{J} \setminus j} u_{ik'}] = \log\left(\sum_{k \in \mathcal{J}} \exp[V_{ik}]\right) - \log\left(\sum_{k \in \mathcal{J} \setminus j} \exp[V_{ik}]\right)$$

Exploit the following:

1. the individual outside good choice probability $s_{i0}(\mathcal{J}, x) = \frac{1}{\sum_{k \in \mathcal{J}} \exp[V_{ik}]}$;
2. that the outside good choice probability after removing j increases by the individual share of j times the individual diversion ratio from j to the outside good $s_{i0}(\mathcal{J} \setminus j, x) = s_{i0}(\mathcal{J}, x) + D_{j0,i}(x) \cdot s_{ij}(x)$;
3. for members of the logit family: $D_{j0,i}(x) = \frac{s_{i0}(x)}{1 - s_{ij}(x)}$.

Welfare and WTP

To get EV we integrate over $\frac{1}{|\alpha_i|}$.

$$\begin{aligned} WTP_i(j) &= \log\left(\frac{s_{i0}(\mathcal{J}\setminus j, x)}{s_{i0}(\mathcal{J}, x)}\right) = \log\left(1 + \frac{D_{j0,i}(x) s_{ij}(x)}{s_{i0}(\mathcal{J}, x)}\right) \\ &= \log\left(1 + \frac{s_{i0}(\mathcal{J}, x)\cdot s_{ij}(x)}{(1 - s_{ij}(x))\cdot s_{i0}(\mathcal{J}, x)}\right) \\ &= \log\left(1 + \frac{s_{ij}(x)}{1 - s_{ij}(x)}\right) \approx \frac{s_{ij}(x)}{1 - s_{ij}(x)} \end{aligned}$$

Individual willingness to pay is exclusively about own share (not presence of close substitutes). Is this a good property or not?

Practical Tips: Checklist

- How much diversion to outside good?
- What are top substitutes by diversion? How much do they get?
- Do identities of best substitutes differ across products?
- How much diversion goes to own products vs. competitors?
- For mergers, what would divestiture candidates be?