30 Years of BLP...

Chris Conlon

IIOC 2024

NYU Stern School of Business and NBER

Thank Organizers+Steve

Thank Coauthors

- w/ Julie Mortimer
    - *Demand Estimation Under Incomplete Product Availability*
    - *Empirical Properties of Diversion Ratios*
    - w/ Paul Sarkis *Estimating Preferences and Substitution Patterns from Second-Choice Data Alone*
- w/ Nirupama Rao
    - *The Cost of Curbing Externalities with Market Power: Alcohol Regulations and Tax Alternatives*
- w/ Matt Backus and Michael Sinkinson
    - *Common Ownership and Competition in the Ready-To-Eat Cereal Industry*
- w/ Jeff Gortmaker
    - *Best Practices for Demand Estimation with pyBLP*
    - *Incorporating Micro Data into Differentiated Products Demand Estimation with PyBLP*
    - *Common Ownership and Competition in the Ready-To-Eat Cereal Industry*

Organization

Past Lots of good ideas in "original" BLP95/99; including some key ones that got ignored and hopefully rediscovered.

- BLP was at its core about simultaneous supply and demand.
- Much like Chamberlain (1987) the section on optimal IV was ahead of its time.

Present Data and Computers are way better than in 1995

- Especially Micro Data/ Mini case study
- Trying to cram everything in the BLP box is not always the best idea.
  - Some BLP alternatives: analytic inverses, approximations, etc.

Future Can we realize the dream of non-parametric identification in estimation?

- Doing ML not just measuring its impact!

The "Classics": Thing #1 Supply Side

Supply Side

Consider the multi-product Bertrand FOCs where $\arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p})$:

$$\pi_f(\mathbf{p}) \equiv \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot s_j(\mathbf{p}) + \sum_{k \in \mathcal{J}_g} (p_k - c_k) \cdot s_k(\mathbf{p})$$

$$\rightarrow 0 = s_j(\mathbf{p}) + \sum_{k \in \mathcal{J}_f} (p_k - c_k) \frac{\partial s_k}{\partial p_j}(\mathbf{p})$$

It is helpful to define the cross derviative matrix $\Delta_{(j,k)}(\mathbf{p}) = -\frac{\partial s_j}{\partial p_k}(\mathbf{p})$, and the ownership matrix:

$$\mathcal{H}_{(j,k)} = \left\{ \begin{array}{lr} 1 & \text{for } (j,k) \in \mathcal{J}_f \text{ for any } f \\ 0 & \text{o.w} \end{array} \right\}$$

We can re-write the FOC in matrix form where $\odot$ denotes Hadamard product (element-wise):

$$\begin{aligned} s(\mathbf{p}) &= (\mathcal{H} \odot \Delta(\mathbf{p})) \cdot (\mathbf{p} - \mathbf{mc}), \\ \mathbf{p} - \mathbf{mc} &= \underbrace{(\mathcal{H} \odot \Delta(\mathbf{p}))^{-1} s(\mathbf{p})}_{\eta(\mathbf{p}, \mathbf{s}, \theta_2)}. \end{aligned}$$

What's the point?

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

Demand systems have two main deliverables:

- Own-price elasticities $\epsilon_{jj}(\mathbf{p})$
- Substitution patterns
  - Cross elasticities: $\epsilon_{jk}(\mathbf{p}) = \frac{\partial q_k}{\partial p_j}$
  - Diversion Ratios: $D_{jk}(\mathbf{p}) = \frac{\partial q_k}{\partial p_j} / |\frac{\partial q_j}{\partial p_j}|$
- Other checks: $D_{j0}(\mathbf{p})$ diversion to outside good; ϵ^{agg} category elasticity to 1% tax.

We did Nash-in-Prices because it is popular but we could have done something else.

What keeps me up at night?

- The unrestricted matrix of of $D_{jk}(\mathbf{p})$'s (or elasticities) is $J \times J$ and probably not feasible to estimate directly. $\rightarrow$ need some dimension reduction.
- Logit simply assumes proportional substitution $D_{jk} = \frac{s_k}{1-s_j}$ so that $\mathbf{D}(\mathbf{p})$ is of rank one!
- The BLP solution is to project $\mathbf{D}(\mathbf{p})$ onto a lower-dimensional basis of x_j characteristics.
  - Ultimately the basis will be only as good as the characteristics x_{jt} with heterogeneous coefficients.
  - Distributional assumptions on $f(\beta_i)$ (ie: independent normal) further restrict the basis.
- The hardest thing to match is typically substitution to closest substitutes.
- I worry that most BLP models (one RC, etc.) look too much like the plain logit.

Constructing Supply Moments

If we are willing to impose $MR = MC$ (as in the original BLP papers) we can recover implied markups/ marginal costs:

$$\mathbf{mc}(\theta_2) \equiv \mathbf{p} - \boldsymbol{\eta}(\mathcal{S}_t, \mathbf{p}_t, \chi_t, y_t, ; \theta_2)$$
$$f\left(\mathbf{p} - \boldsymbol{\eta}(\mathcal{S}_t, \mathbf{p}_t, \chi_t, y_t, ; \theta_2)\right) = [\mathrm{x}_{jt}\ , \mathrm{w}_{jt}]\theta_3 + g(s_{jt}) + \omega_{jt}$$

- $f(\cdot)$ is usually $\log(\cdot)$ or identity; it is actually a production function
- $g(s_{jt})$ captures returns to scale and requires an additional instrument

Simultaneous Supply and Demand

$$\sigma_j^{-1}(\mathcal{S}_t, \mathbf{p}_t, \chi_t, y_t; \theta_2) = [\mathrm{x}_{jt}, \mathrm{v}_{jt}]\,\theta_1 - \alpha p_{jt} + \xi_{jt}$$
$$f\left(p_{jt} - \eta_{jt}(\mathcal{S}_t, \mathbf{p}_t, \chi_t, y_t; \theta_2)\right) = [\mathrm{x}_{jt}, \mathrm{w}_{jt}]\,\theta_3 + \omega_{jt}$$

We can now form two sets of moments: $\mathbb{E}[\omega_{jt} \mid z_{jt}^s] = 0$ and $\mathbb{E}[\xi_{jt} \mid z_{jt}^d] = 0$

- These provide overidentifying restrictions for (θ_2, α)
- Conditional on θ_2 (distribution of random coefficients) and α this is just linear IV-GMM again.
- The derivatives $\left(\frac{\partial \xi_{jt}}{\partial \theta_2}, \frac{\partial \omega_{jt}}{\partial \theta_2}\right)$ beacuse of $\frac{\partial \eta_{jt}}{\partial \theta_2}$ in particular, are complicated (But `PyBLP` knows how to do these).
- As Steve has made clear this is likely a many weak IV situation many potential IV's (others $x_{-j}, w_{-j}, v_{-j}, y_t$), but hard to know which are strong.

The "Classics": Thing #2
Optimal Instruments

Optimal Instruments (Chamberlain 1987)

Chamberlain (1987) asks how can we choose $f(z_i)$ to obtain the semi-parametric efficiency bound with conditional moment restrictions:

$$\mathbb{E}[g(z_i,\theta)|z_i] = 0 \Rightarrow \mathbb{E}[g(z_i,\theta) \cdot f(z_i)] = 0$$

Recall that the asymptotic GMM variance depends on $(G' \, \Omega^{-1} G \,)$

The answer is to choose instruments related to the (expected) Jacobian of moment conditions w.r.t θ. The true Jacobian at θ_0 is infeasible:

$$G = \mathbb{E}\left[\frac{\partial g(z_i,\theta)}{\partial \theta} | z_i, \theta_0\right]$$

Problems: we don't know θ_0 and endogeneity.

Chamberlain (1987)

Chamberlain (1987) showed that the approximation to the optimal instruments are given by the expected Jacobian contribution for each observation (j,t): $\mathbb{E}[G_{jt}(\mathbf{Z_t})\,\Omega_{jt}^{-1}|\mathbf{Z_t}]$. For BLP this amounts to:

$$G = \mathbb{E}\left[\left(\frac{\partial \xi_{jt}}{\partial \theta}, \frac{\partial \omega_{jt}}{\partial \theta}\right)|\mathbf{Z_t}\right], \quad \Omega = \mathbb{E}\left[\begin{pmatrix}\xi_{jt}\\ \omega_{jt}\end{pmatrix}\begin{pmatrix}\xi_{jt} & \omega_{jt}\end{pmatrix}|\mathbf{Z_t}\right]$$

$$\begin{aligned}\xi_{jt} &= \sigma_j^{-1}(\cdot,\theta_2) - [\mathrm{x}_{jt}\,,\mathrm{v}_{jt}]\,\theta_1 + \alpha p_{jt}\\ \omega_{jt} &= f\left(p_{jt} - \eta_{jt}(\cdot,\theta_2)\right) - [\mathrm{x}_{jt}\,,\mathrm{w}_{jt}]\,\theta_3 + \omega_{jt}\end{aligned}$$

For the exogenous variables: $\mathbb{E}\left[\frac{\partial \xi_{jt}}{\partial \theta_1} \mid z_{jt}^d\right] = [\mathrm{x}_{jt}\,,\mathrm{v}_{jt}]$ and $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \theta_3} \mid z_{jt}^s\right] = [\mathrm{x}_{jt}\,,\mathrm{w}_{jt}]$.

For the endogenous prices: $\mathbb{E}\left[\frac{\partial \xi_{jt}}{\partial \alpha} \mid z_{jt}^d\right] = \mathbb{E}[p_{jt} \mid z_{jt}^d]$ and $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \alpha} \mid z_{jt}^s\right] = \mathbb{E}[f'(\cdot)(p_{jt} - \frac{\partial \eta_{jt}}{\partial \alpha}) \mid z_{jt}^s]$.

For the endogenous θ_2: $\mathbb{E}\left[\frac{\partial \xi_{jt}}{\partial \theta_2} \mid z_{jt}^d\right] = \mathbb{E}\left[\left[\frac{\partial \sigma_{jt}}{\partial \xi_{jt}}\right]^{-1}\left[\frac{\partial \sigma_{jt}}{\partial \theta_2}\right] \mid z_{jt}^d\right]$ and $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \theta_2} \mid z_{jt}^s\right]$

(but you can't condition on p_{jt})

Optimal Instruments

Even with an intitial guess of $\widehat{\theta}$, we still have that p_{jt} or η_{jt} depends on (ω_j, ξ_t) in a highly nonlinear way (no explicit solution!). But we have some options:

- Pray to the God of Sieves:
  - Since any $f(x, z)$ satisfies our orthogonality condition, we can try to choose $f(x, z)$ as a basis to approximate optimal instruments. (Newey 1990)
  - This is challenging in practice – and in fact suffers from a curse of dimensionality.
  - This is frequently given as a rationale behind higher order x's.
- Plug in a guess for first stage p_{jt}:
  - Reynaert Verboven (2014) suggest $\mathbb{E}[p_{jt} \mid \mathrm{x}_{jt}, \mathrm{w}_{jt}] = mc_{jt}$ (perfect competition), but might as well include other z_{jt}^d (like BLP instruments).
  - $\mathbb{E}[p_{jt} \mid z_{jt}^d]$ is easy, and non-parametric regression is pretty good.
- Use the nonlinearity in the model! (BLP 199)

Feasible Recipe (BLP 1999)

1. Fix $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3)$ and draw $(\boldsymbol{\xi}^*, \boldsymbol{\omega}^*)$ from empirical density
2. Solve firm FOC's for $\hat{\mathbf{p}}_\mathbf{t}(\boldsymbol{\xi}^*, \boldsymbol{\omega}^*, \widehat{\theta})$ and shares $\mathbf{s}_\mathbf{t}(\hat{\mathbf{p}}_\mathbf{t}, \widehat{\theta})$
3. Compute necessary Jacobian
4. Average over multiple values of$(\boldsymbol{\xi}^*, \boldsymbol{\omega}^*)$. (Lazy approach: use only $(\boldsymbol{\xi}^*, \boldsymbol{\omega}^*) = 0$).

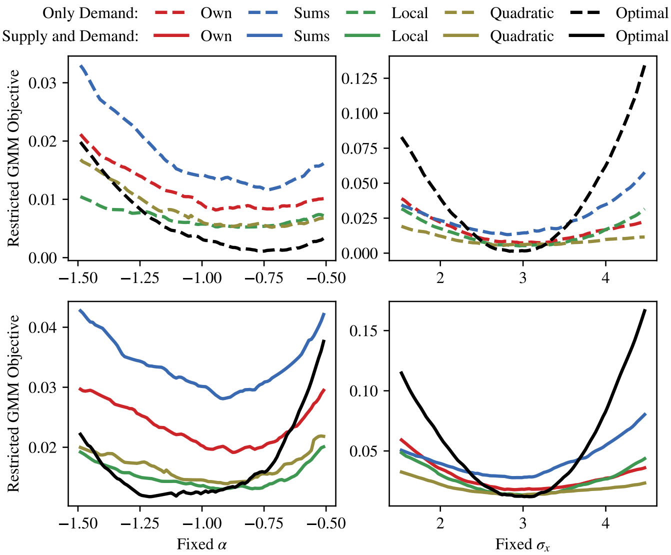
In simulation the "lazy" approach does just as well.

Alternative: Can we use $\mathbb{E}[\mathbf{p}_\mathbf{t} \mid \mathbf{Z}_\mathbf{t}]$ instead for (2) if we don't have supply side

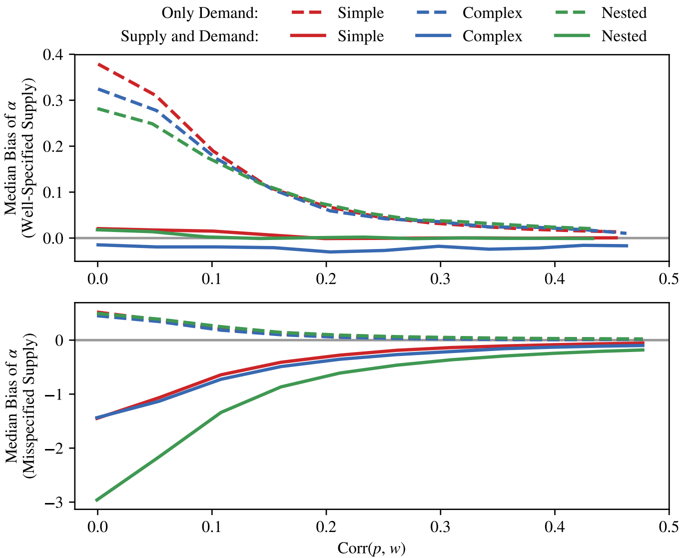
IV Comparison: Conlon and Gortmaker (2020)

				True Value				Median Bias				Median Absolute Error			
Simulation	Supply	Instruments	Seconds	α	σ_x	σ_p	ρ	α	σ_x	σ_p	ρ	α	σ_x	σ_p	ρ
Simple	No	Own	0.6	-1	3			0.126	-0.045			0.238	0.257		
Simple	No	Sums	0.6	-1	3			0.224	-0.076			0.257	0.208		
Simple	No	Local	0.6	-1	3			0.181	-0.056			0.242	0.235		
Simple	No	Quadratic	0.6	-1	3			0.206	-0.085			0.263	0.239		
Simple	No	Optimal	0.8	-1	3			0.218	-0.049			0.250	0.174		
Simple	Yes	Own	1.4	-1	3			0.021	0.006			0.226	0.250		
Simple	Yes	Sums	1.5	-1	3			0.054	-0.020			0.193	0.196		
Simple	Yes	Local	1.4	-1	3			0.035	-0.006			0.207	0.229		
Simple	Yes	Quadratic	1.4	-1	3			0.047	-0.022			0.217	0.237		
Simple	Yes	Optimal	2.2	-1	3			0.005	0.012			0.170	0.171		
Complex	No	Own	1.1	-1	3	0.2		-0.025	0.000	-0.200		0.381	0.272	0.200	
Complex	No	Sums	1.1	-1	3	0.2		0.225	-0.132	-0.057		0.263	0.217	0.200	
Complex	No	Local	1.0	-1	3	0.2		0.184	-0.107	-0.085		0.274	0.236	0.200	
Complex	No	Quadratic	1.0	-1	3	0.2		0.200	-0.117	-0.198		0.299	0.243	0.200	
Complex	No	Optimal	1.6	-1	3	0.2		0.191	-0.119	0.001		0.274	0.195	0.200	
Complex	Yes	Own	3.9	-1	3	0.2		-0.213	0.060	0.208		0.325	0.263	0.208	
Complex	Yes	Sums	3.3	-1	3	0.2		0.018	-0.104	0.052		0.203	0.207	0.180	
Complex	Yes	Local	3.4	-1	3	0.2		-0.043	-0.078	0.135		0.216	0.225	0.200	
Complex	Yes	Quadratic	3.5	-1	3	0.2		-0.028	-0.067	0.116		0.237	0.227	0.200	
Complex	Yes	Optimal	4.9	-1	3	0.2		-0.024	-0.036	-0.002		0.193	0.171	0.191	

IV Comparison: Conlon and Gortmaker (2020)



Cost Shifters Really Matter (from Conlon Gortmaker RJE)



Aside: Optimal IV Everywhere! (Backus, Conlon, Sinkinson)

In our paper on testing conduct we are interested in testing $H_0 : \tau = 1$ and $H_a : \tau_= 0$

$$\omega_{jt} = p_{jt} - \tau \cdot \eta_{jt}^{(a)} - (1 - \tau) \cdot \eta_{jt}^{(b)} - h(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3)$$

The key to the test is to realize that optimal IV: $\mathbb{E}\left[\frac{\partial \omega_{jt}}{\partial \tau} \mid z_{jt}^s\right] = \mathbb{E}\left[\eta_{jt}^{(a)} - \eta_{jt}^{(b)} \mid z_{jt}^s\right]$.

- Instruments predict the difference in markups!
- Can run one non-parametric regression for $\mathbb{E}\left[\eta_{jt}^{(a)} - \eta_{jt}^{(b)} \mid z_{jt}^s\right]$ and another for the nuisance function (observed markup shifters) $h(\cdot)$.
- This is an easy way to do Berry Haile (2014). Duarte, Lorenzo Magnolfi, Mikkel Sølvsten, Christopher Sullivan (2024) get a similar expression with a different approach.

Takeaway

What does this mean:

- Optimal IV aren't magic, you probably need good cost shifters.
- We should always check $\mathbb{E}[\mathbf{p} \mid \mathbf{z}]$ before we do anything else.
- May want to consider adding a supply side (if you're willing to assume for counterfactuals, why not?)
- Certainly should do `results.compute_optimal_instruments()` in PyBLP.

Adding Micro Data

Explain Micro Data Here

These slides are broken

Micro Data Gotchas

Alternative: GRMPS

Quick Case Study: Supply and Micro Data

Supply and Micro Data: (Conlon Rao 2023)

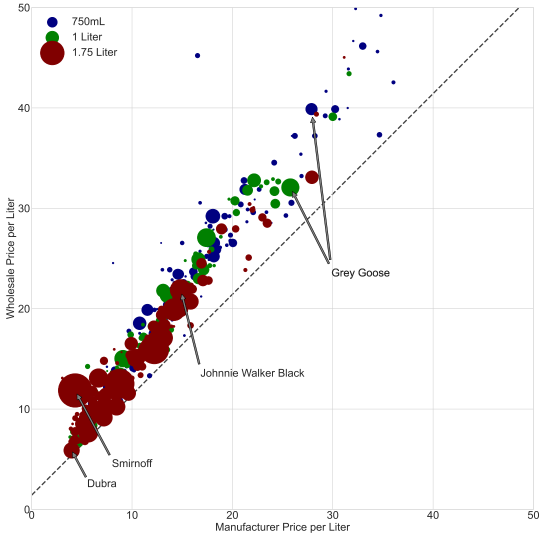
Consumer i chooses product j (brand-size-flavor) in quarter t:

$$u_{ijt} = \beta_i^0 - \alpha_i \, p_{jt} + \beta_i^{1750} \cdot \mathbb{I}[1750mL]_j + \gamma_j + \gamma_t + \varepsilon_{ijt}(\rho)$$

$$\begin{pmatrix} \ln \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \overline{\alpha} \\ \theta_1 \end{pmatrix} + \Sigma \cdot \nu_i + \sum_k \Pi_k \cdot \mathbb{I}\{LB_k \leqslant \text{Income}_i < UB_k\}$$

- Nesting Parameter ρ: Substitution within category (Vodka, Gin, etc.)
- Consumers of different income levels have different mean values for coefficients
- Conditional on income, normally distributed unobserved heterogeneity for:
  - Price α_i
  - Constant β_i^0 (Overall demand for spirits)
  - Package Size: β_i^{1750} (Large vs. small bottles)

Wholesale Margins Under Post and Hold



- Price Cost Margins (and Lerner Markups) are higher on premium products
- Markups on least expensive products (plastic bottle vodka) are very low.
- Smirnoff (1.75L) is best seller (high markup / outlier).
- A planner seeking to minimize ethanol consumption would flatten these markups!
- Matching this pattern is kind of the whole ballgame !
- Plain logit gives $\epsilon_{jj} = \alpha \cdot p_j \cdot (1 - s_j)$.

Demand Estimates (from PyBLP, Conlon Gortmaker (2020, 2023))

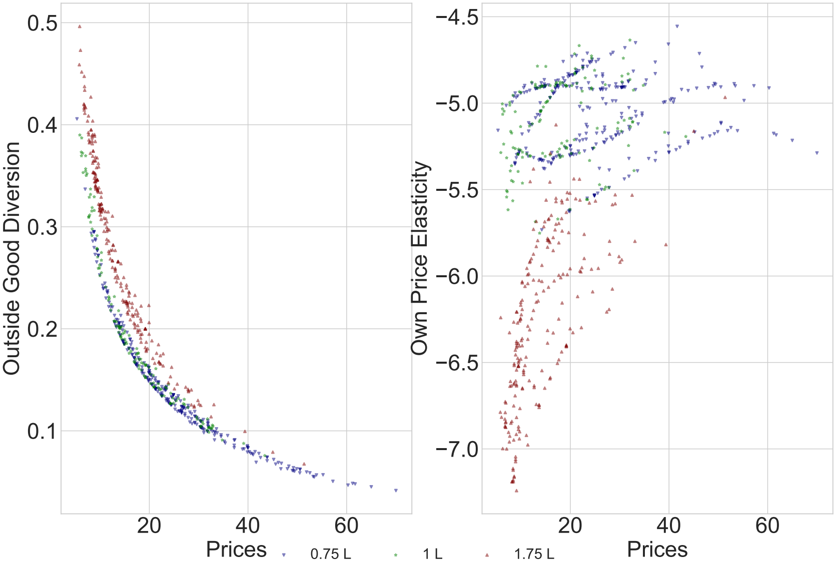
Π	Const	Price	1750mL
Below \$25k	2.928	-0.260	0.543
	(0.233)	(0.056)	(0.075)
\$25k-\$45k	0.184	-0.170	0.536
	(0.236)	(0.054)	(0.083)
\$45k-\$70k	0.000	-0.179	0.980
	(0.000)	(0.053)	(0.093)
\$70k-\$100k	-0.452	-0.496	0.608
	(0.227)	(0.051)	(0.079)
Above \$100k	-1.777	-1.543	0.145
	(0.234)	(0.047)	(0.055)
Σ^2			
Price	0.000	0.697	0.695
	(0.107)	(0.028)	(0.048)
1750mL	0.000	0.695	1.167
	(0.086)	(0.048)	(0.236)
Nesting Parameter ρ		0.423	
		(0.026)	
Fixed Effects		Brand+Quarter	
Model Predictions	25%	50%	75%
Own Elasticity: $\frac{\partial \log q_j}{\partial \log p_j}$	-5.839	-5.162	-4.733
Aggregate Elasticity: $\frac{\partial \log Q}{\partial \log P}$	-0.333	-0.329	-0.322
Own Pass-Through: $\frac{\partial p_j}{\partial c_j}$	1.256	1.284	1.320
Observed Wholesale Markup (PH)	0.188	0.233	0.276
Predicted Wholesale Markup (PH)	0.205	0.231	0.259

- Demographic Interactions w/ 5 income bins (matched to micro-moments)
- Correlated Normal Tastes: (Constant, Large Size, Price)
- Supply moments exploit observed upstream prices and tax change (ie: match observed markups).

$$\mathbb{E}[\omega_{jt}] = 0, \text{ with } \omega_{jt} = \left(p_{jt}^{w} - p_{jt}^{m} - \tau_{jt}\right) - \eta_{jt}\left(\theta_2\right).$$

- Match estimate of aggregate elasticity from tax change $\varepsilon = -0.4$.
- Pass-through consistent with estimates from our AEJ:Policy paper.

Elasticities and Diversion Ratios



Diversion Ratios

	Median Price	% Substitution		Median Price	% Substitution
Capt Morgan Spiced 1.75 L ($15.85)			Cuervo Gold 1.75 L ($18.33)		
Bacardi Superior Lt Dry Rum 1.75 L	12.52	13.07	Don Julio Silver 1.75 L	22.81	5.00
Bacardi Dark Rum 1.75 L	12.52	2.71	Cuervo Gold 1.0 L	21.32	3.82
Bacardi Superior Lt Dry Rum 1.0 L	15.03	2.44	Sauza Giro Tequila Gold 1.0 L	8.83	3.07
Smirnoff 1.75 L	11.85	2.36	Smirnoff 1.75 L	11.85	2.44
Lady Bligh Spiced V Island Rum 1.75 L	9.43	2.18	Absolut Vodka 1.75 L	15.94	2.06
Woodford 0.75 L ($34.55)			Beefeater Gin 1.75 L ($17.09)		
Jack Daniel Black Label 1.0 L	27.08	7.66	Tanqueray 1.75 L	17.09	12.80
Jack Daniel Black Label 1.75 L	21.85	4.91	Gordons 1.75 L	11.19	4.14
Jack Daniel Black Label 0.75 L	29.21	4.83	Seagrams Gin 1.75 L	10.23	2.85
Makers Mark 1.0 L	32.79	4.52	Bombay 1.75 L	21.95	2.27
Makers Mark 0.75 L	31.88	2.80	Smirnoff 1.75 L	11.85	2.27
Dubra Vdk Dom 80P 1.75 L ($5.88)			Belvedere Vodka 0.75 L ($30.55)		
Popov Vodka 1.75 L	7.66	7.56	Grey Goose 1.0 L	32.08	5.09
Smirnoff 1.75 L	11.85	3.15	Absolut Vodka 1.75 L	15.94	3.82
Sobieski Poland 1.75 L	9.09	3.14	Absolut Vodka 1.0 L	24.91	2.74
Grays Peak Vdk Dom 1.75 L	9.16	2.87	Smirnoff 1.75 L	11.85	2.43
Wolfschmidt 1.75 L	6.92	2.48	Grey Goose 0.75 L	39.88	2.22

Unused: Exclusion Restrictions (see Berry Haile 2014)

$$\begin{aligned} \delta_{jt}(\mathcal{S}_t, \mathrm{y}_t, \widetilde{\theta}_2) &= [\mathrm{x}_{jt}, \mathrm{v}_{jt}]\beta - \alpha p_{jt} + \xi_{jt} \\ f(p_{jt} - \eta_{jt}(\theta_2, \mathbf{p}, \mathbf{s})) &= h(\mathrm{x}_{jt}, \mathrm{w}_{jt}; \theta_3) + \omega_{jt} \end{aligned}$$

The first place to look for exclusion restrictions/instruments:

- Something in another equation!
- v_j shifts demand but not supply
- w_j shifts supply but not demand
- y_t is a sneaky demand shifter
- If it doesn't shift either is it really relevant?

Alternative: MacKay Miller (2022) propose $Cov(\xi_{jt}, \omega_{jt}) = 0$ as an alternative.

Intuition from Linear IV (FRAC: Salanie and Wolak)

Simple case where $\theta_0 = (\beta_0, \pi_0, \sigma_0)'$. A second-order Taylor expansion around $\pi_0 = \sigma_0 = 0$ gives the following linear model with four regressors:

$$\log \frac{\mathcal{S}_{jt}}{\mathcal{S}_{0t}} \approx \beta_0 x_{jt} + \sigma_0^2 a_{jt} + \pi_0 m_t^y x_{jt} + \pi_0^2 v_t^y a_{jt} + \xi_{jt}, \quad a_{jt} = \left(\frac{x_{jt}}{2} - \sum_{k \in \mathcal{J}_t} \mathcal{S}_{kt} \cdot x_{kt} \right) \cdot x_{jt} \tag{1}$$

- $m_t^y = \sum_{i \in \mathcal{I}_t} w_{it} \cdot y_{it}$ is the within-market demographic mean
- $v_t^y = \sum_{i \in \mathcal{I}_t} w_{it} \cdot (y_{it} - m_t^y)^2$ is its variance
- a_{jt} is an "artificial regressor" that reflects within-market differentiation of the product characteristic x_{jt}.
- Linear but we still need an IV for a_{jt}.

Implemented in Julia by Jimbo Brand `https://github.com/jamesbrandecon/FRAC.jl`

Connection or when do GH IV work well?

Recall the GH IV are:

$$J \cdot x_{jt}^2 + \underbrace{\sum_k x_{kt}^2}_{\text{constant for } t} - 2\sum_k x_{jt} \cdot x_{kt}$$

and the artificial regressor is

$$\frac{1}{2} x_{jt}^2 - 2 x_{jt} \cdot \sum_k S_{kt} \cdot x_{kt}$$

- We should be share weighting the interaction term, but GH assume equal weighting.
- Should be able to do better than these IV (but ideal is infeasible...)
- Alternative take: GH propose IIA test that looks a lot like Salanie Wolak estimator. Good for starting values? Or as pre-test for heterogeneity?
- Warning: I find these are always nearly colinear and run PCA first...

Future Stuff

Embeddings: Magnolfi Maclure Sorensen 1

Embeddings: Magnolfi Maclure Sorensen 2

Conlon Mortimer Sarkis 1

Conlon Mortimer Sarkis 2

Conlon Mortimer Sarkis 3

Other nonparametrics?

Thanks!