Grad IO Lecture 10

Lorenzo Magnolfi

Department of Economics, UW-Madison

Thursday 8th October, 2020

Back to Berry (1994)

- Suppose we have product-market-level data (p_{jm}, x_{jm}, S_{jm}) for products $j = 1, ..., J$ and markets $m = 1, ..., M$
- Suppress m in what follows for simplicity
- Suppose we model discrete choice demand as arising from a ***random utility model***
  - Agents have indirect utility u_{ij} for each good, choose $j_i^* = \arg\max_j u_{ij}$
- Want to take seriously presence of unobserved product characteristics

Berry (1994) - Model

- Consider utility specification:

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + \varepsilon_{ij},$$

  no random coeffs, *unobservable product characteristic* ξ_j

  - only difference with MNL!

- Can write $\delta_j = x_j\beta - \alpha p_j + \xi_j$, and expression for market shares is:

$$s_j(\delta) = \frac{e^{\delta_j}}{\sum_k e^{\delta_k}},$$

  and adopt normalization $\delta_0 = 0$

Berry (1994) - Inversion

- Notice that

$$\ln \frac{S_j}{S_0} = \delta_j,$$

so immediate analytical way to perform the **inversion** in this case and obtain $\delta = s^{-1}(S)$

- Since $\delta_j = x_j\beta - \alpha p_j + \xi_j$, we have

$$\ln \frac{S_j}{S_0} = x_j\beta - \alpha p_j + \xi_j.$$

- Coefficients are identified as long as we have sufficient number of instruments z correlated with (x, p) and such that $E(\xi|z) = 0$

Berry (1994) - The Supply Side

- Start from simple linear specification for marginal costs

$$mc_j = x_j^{Cost}\eta + \lambda q_j + \omega_j$$

- Differentiated products framework allows for presence of markups, e.g. Bertrand price setting equilibrium is

$$\begin{aligned} p_j &= mc_j + markup_j, \\ &= x_j^{Cost}\eta + \lambda q_j + \frac{q_j}{|\frac{\partial q_j}{\partial p_j}|} + \omega_j, \end{aligned}$$

  and $|\frac{\partial q_j}{\partial p_j}|$ is known if the demand system has already been estimated
  - Need instrument for q (demand shifter)

Berry (1994) Remarks/Questions

Key steps for **identification**:

- Inversion of market shares to recover

$$\delta(S) = x_j \overline{\beta} - \alpha p_j + \xi_j,$$

- Instruments z such that $E(\xi|z) = 0$

Then need to come up with *feasible* ***estimation*** *strategy*!

- Why cannot have ξ_j as FE?
- Why not form a Likelihood function?
- Why not just use the "reduced form" of the model?

Berry (1994) - Substitution Patterns?

- I presented Berry (1994) in the context of simple MNL demand
  - Introduced the product-specific unobservable ξ, but haven't fixed IIA!
- To fix (alleviate) IIA, introduce *random coefficients*
  - *"Mixed Logit"*
- Consider

$$u_{ijt} = x_{jt}'\beta_i - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt},$$

for

$$\beta_i^l = \overline{\beta} + \sum_l x_j^l \sigma_l \zeta_i^l ; \alpha_i = \alpha + p_j \sigma_p \zeta_i^p,$$

so that can group

$$\delta_j = x_j \beta - \alpha p_j + \xi_j,$$

and model becomes

$$s_j(\delta, \sigma) = \int \frac{e^{\delta_j + \sum_l x_j^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p}}{\sum_k e^{\delta_k + \sum_l x_k^l \sigma_l \zeta_i^l - p_k \sigma_p \zeta_i^p}} dF(\zeta_i)$$

Random Coefficients Fix Substitution Patterns

- Error term $v_{ij} = \sum_l x_j^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p + \varepsilon_{ij}$ is no longer iid, depends on characteristics x
- Consumers who have high v_{ij} are likely to have high v_{ik} for k similar to j in characteristics
- Derivatives are now

$$\frac{\partial s_j}{\partial p_k} = \int_i \alpha s_{ij} s_{ik} dF(\zeta_i)$$

where $s_{ij} = \frac{e^{\delta_j + \sum_l x_j^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p}}{\sum_k e^{\delta_k + \sum_l x_k^l \sigma_l \zeta_i^l - p_k \sigma_p \zeta_i^p}}$

How to Invert Mixed Logit Models?

- Mixed Logit model is now

$$s_j(\delta,\sigma) = \int \frac{e^{\delta_j + \sum_l x_j^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p}}{\sum_k e^{\delta_k + \sum_l x_k^l \sigma_l \zeta_i^l - p_k \sigma_p \zeta_i^p}} dF(\zeta_i),$$

  and we know (Berry, 1994) we need to invert it to employ linear identification strategy

- But analytical inversion available with Logit doesn't work anymore!
- For model with unobs prod char ξ and random coefficients, how to
  - perform inversion *in practice?*
  - implement feasible estimation strategy?
- This is what Berry, Levinsohn and Pakes (1995) is about

Berry, Levinsohn, Pakes (1995)

- **BLP** (1995) implement random coefficient ("mixed Logit") version of model in Berry (1994), and apply it to dataset on car market
  - They develop algorithm to invert market shares (*Nested Fixed Point* - NFP)
  - Devise influential instrumenting strategy ("*BLP Instruments*")
  - This results in a computable GMM function of the parameters, can be used for estimation

Aside on Estimation: GMM

- Sometimes writing H (or equivalently the density h) is impossible, or too expensive, or requires parametric assumptions
- Can define model by just focusing of moments of H
- Consider $m(y,x;\theta)$ such that:

$$\begin{aligned} E\left[m(y,x;\theta_0)\,|x\right] &= \int m(y,x;\theta_0)\,dG(y|x) \\ &= 0\; a.s. - x \end{aligned}$$

- Define moment function $g(\theta;x) = \int m(y,x;\theta) dH(y|x;\theta)$
- Identification: for square, positive sem def W, have

$$\theta \neq \theta_0 \Rightarrow Wg(\theta;x) \neq 0\; a.s. - x$$

- $Q^{GMM}(\theta) = -E^x\left[g(\theta;x)' Wg(\theta;x)\right]$, and if the model is identified, this function has unique maximum at $\theta = \theta_0$

GMM Estimation

- Given a finite sample $\{y_i, x_i\}_{i=1}^n$, we can construct a sample equivalent of Q^{GMM}
- Define objects $g_n(\theta) = \frac{1}{n}\sum_{i=1}^n m(y_i, x_i; \theta)$ and $W_n \to_p W$
- Then, $Q_n^{GMM} = -g_n(\theta)' W_n g_n(\theta)$, and GMM estimator is

$$\hat{\theta}^{GMM} = \arg\max_{\theta \in \Theta} Q_n^{GMM}$$

- Computing GMM estimator only requires evaluating function m!
- (even if analytic form of m is not tractable, can simulate m...)

BLP's GMM strategy

- How to estimate complicated nonlinear model? GMM
- BLP propose GMM objective function for estimation mimics moment conditions $E(\xi|z) = 0$, where z are x and instruments for p (and to pin down σ!)
- Then, let

$$Q_n^{GMM}(\theta) = -\xi(\theta)' zW^{-1}z'\xi(\theta),$$
$$\hat{\theta} = \arg\max_{\theta} Q_n^{GMM}(\theta)$$

where $\xi(\theta) = \delta(\theta) - x'_{jt}\beta + \alpha p_{jt}$ is value of ξ implied by model and θ; W is consistent estimate of $E(z'\xi\xi'z)$

How to Find $\xi(\theta)$?

- Finding $\xi(\theta)$ is same as finding $\delta(\theta)$: *for fixed values of the data (S, x, p), and for fixed θ, bijective mapping between δ and ξ*
- Hence, we are back to the inversion problem of finding $\delta(\theta) = s^{-1}(S; \theta)$
- Recall model

$$s_j(\delta(\theta), \sigma) = \int \frac{e^{\delta_j + \sum_l x_j^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p}}{\sum_k e^{\delta_k + \sum_l x_k^l \sigma_l \zeta_i^l - p_j \sigma_p \zeta_i^p}} dF(\zeta_i),$$

- Draw R values $(\zeta^r)_{r=1,\ldots,R}$ from $F(\cdot)$ and compute

$$s_j(\delta(\theta), \sigma)^R = \frac{1}{R} \sum_{r=1}^{R} \frac{e^{\delta_j(\theta) + \sum_l x_j^l \sigma_l \zeta_i^{l,r} - p_j \sigma_p \zeta_i^{p,r}}}{\sum_k e^{\delta_k(\theta) + \sum_l x_k^l \sigma_l \zeta_i^{l,r} - p_k \sigma_p \zeta_i^{p,r}}}$$

  - Can alternatively extract nonrandom values in way that minimizes $||s - s^R||$ - look up importance sampling, Halton sets

How to compute function s^{-1}?

- Having simulated model mapping, we can set it equal to data

$$s_j(\delta(\theta), \sigma)^R = S_j$$

- Then, finding $s^{-1}(S;\theta)^R$ amounts to finding δ that solves

$$S_j = \frac{1}{R}\sum_{r=1}^{R} \frac{e^{\delta_j + \sum_l x_j^l \sigma_l \zeta_i^{l,r} - p_j \sigma_p \zeta_i^{p,r}}}{\sum_k e^{\delta_k + \sum_l x_k^l \sigma_l \zeta_i^{l,r} - p_k \sigma_p \zeta_i^{p,r}}}$$

  for fixed θ, S

- BLP propose an algorithm for finding such δ

Computation of BLP: NFXP

- Step 0: fix θ
- Step 1: starting from arbitrary δ^0, iterate the contraction

$$\delta_j^k(\theta) = \delta_j^{k-1}(\theta) + \log S_j - \log s_j\left(\delta^{k-1}(\theta), \theta\right)^R$$

  until convergence to $\delta^K(\theta)$ i.e. $||\delta^K(\theta) - \delta^{K-1}(\theta)|| \leq tol$ for some small tolerance value, where s^R indicates that the market share is simulated
- Step 2: use $\delta^K(\theta)$ to compute objective function $Q_n^{GMM}(\theta)$
- Step 3: continue search over θ until you find a maximum of $Q_n^{GMM}(\theta)$

Some Computational Considerations on NFXP

The algorithm is costly to program and to compute; some issues:

- Using last generation numerical integration techniques when computing simulated integral can significantly increase stability of inner-loop and performance (Skrainka and Judd 2011)
- The outer loop can converge to local maxima, need to use multiple starting points, perhaps global optimization methods (Knittel and Metaxoglu 2014)
- Instruments play a key role in robustness of algorithm

What Instruments?

- We need instruments w to generate moments; at least as many as # of parameters
- x is usually assumed to be exogenous
- Natural instrument for price: supply shifter e.g. cost variable - rarely available
- *BLP instruments*: functions of characteristics of other products that correlate with markups and hence price
  - Think of how markups depend on "location" in the product space in models of differentiated products
  - Armstrong (2015): as the number of products increases, BLP instruments become weaker

What Instruments?

- *Hausman instruments*: prices of same good in other markets
  - But what if correlation is due to common demand shocks (e.g. national ad campaign), as opposed to common cost shocks? (Bresnahan 1997)
- *Waldfogel instruments*: features of the distribution of consumer characteristics in market (why valid?)
- Up to here, focus on instrument for p; which moments pin down σ? most studies are fuzzy on this issue

Optimal Instruments

- Chamberlain (1987): when GMM objective function is

$$Q_n^{GMM}(\theta) = -\xi(\theta)' z W^{-1} z' \xi(\theta),$$

then optimal instrument matrix is

$$z^* = E\left[\frac{\partial \xi(\theta_0)}{\partial \theta'} | z\right]' E\left[\xi \xi' | z\right]$$

- Since θ_0 is unknown parameter value, these can only be approximated
- Reynaert and Verboven (2014): using optimal instruments is key for the efficiency of the estimator and the stability of the algorithm
- Gandhi and Houde (2017): many important lessons, including
  - Be mindful of *weak identification*
  - A practical instrumenting strategy based on differentiation

BLP's Empirical Application

- Big contribution of BLP: proposing a practical method
- Use data on all car models sold in the US in 1971-1990: annual sales, car characteristics (weight, hp, dimensions, MPG,...) and list retail price for base model in 1983 USD
- Use as instruments characteristics of competitor's products (stiffer competition means lower markup) and characteristics of own products (multiproduct firms will price less aggressively not to cannibalize own product)

BLP's Empirical Application

- Use also model of supply side:

$$p = mc + a(p, x, \xi; \theta)$$

where a is a markup term that's fully determined by data (p, x), unobservable characteristics ξ and θ;

- Marginal costs are parametrized as function of observables plus unobservable component

$$mc = \exp\left(x'\gamma + x^{c\prime}\gamma^c + \omega\right),$$

where x^c are characteristics that only enter costs, so that

$$\omega = \log\left(p - a(p, x, \xi; \theta)\right) - x'\gamma + x^{c\prime}\gamma^c,$$

and additional moments based on $E(\omega | x, x^c) = 0$ enter the GMM objective together with demand side moments

BLP's Empirical Application

TABLE IV

ESTIMATED PARAMETERS OF THE DEMAND AND PRICING EQUATIONS: BLP SPECIFICATION, 2217 OBSERVATIONS

Demand Side Parameters	Variable	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
Means ($\bar{\beta}$'s)	*Constant*	−7.061	0.941	−7.304	0.746
	HP/Weight	2.883	2.019	2.185	0.896
	Air	1.521	0.891	0.579	0.632
	MP\$	−0.122	0.320	−0.049	0.164
	Size	3.460	0.610	2.604	0.285
Std. Deviations (σ_β's)	*Constant*	3.612	1.485	2.009	1.017
	HP/Weight	4.628	1.885	1.586	1.186
	Air	1.818	1.695	1.215	1.149
	MP\$	1.050	0.272	0.670	0.168
	Size	2.056	0.585	1.510	0.297
Term on Price (α)	$\ln(y-p)$	43.501	6.427	23.710	4.079
Cost Side Parameters					
	Constant	0.952	0.194	0.726	0.285
	ln(*HP/Weight*)	0.477	0.056	0.313	0.071
	Air	0.619	0.038	0.290	0.052
	ln(*MPG*)	−0.415	0.055	0.293	0.091
	ln(*Size*)	−0.046	0.081	1.499	0.139
	Trend	0.019	0.002	0.026	0.004
	$\ln(q)$			−0.387	0.029

BLP's Empirical Application

TABLE VII

SUBSTITUTION TO THE OUTSIDE GOOD

Model	Given a price increase, the percentage who substitute to the outside good (as a percentage of all who substitute away.) Logit	BLP
Mazda 323	90.870	27.123
Nissan Sentra	90.843	26.133
Ford Escort	90.592	27.996
Chevy Cavalier	90.585	26.389
Honda Accord	90.458	21.839
Ford Taurus	90.566	25.214
Buick Century	90.777	25.402
Nissan Maxima	90.790	21.738
Acura Legend	90.838	20.786
Lincoln Town Car	90.739	20.309
Cadillac Seville	90.860	16.734
Lexus LS400	90.851	10.090
BMW 735*i*	90.883	10.101

- *Ackerberg, Daniel, et al. "Econometric tools for analyzing market outcomes." Handbook of econometrics 6 (2007): 4171-4276.*
- Armstrong, Timothy B. "Large Market Asymptotics for Differentiated Product Demand Estimators with Economic Models of Supply." (2015).
- Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile prices in market equilibrium." *Econometrica* (1995): 841-890.
- Bresnahan, Tim F. "The apple-cinnamon cheerios war: Valuing new goods, identifying market power, and economic measurement." The Economics of New Goods. The University of Chicago Press, Chicago/London (1997).
- Conlon, Christopher, and Jeff Gortmaker. Best practices for differentiated products demand estimation with pyblp. Working paper. url: https://chrisconlon. github. io/site/pyblp. pdf, 2019.

- Dubé, Jean-Pierre, Jeremy T. Fox, and Che-Lin Su. "Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation." Econometrica 80.5 (2012): 2231-2267.
- Hausman, Jerry A. "Valuation of new goods under perfect and imperfect competition." The economics of new goods. University of Chicago Press, 1996. 207-248.
- Knittel, Christopher R., and Konstantinos Metaxoglou. "Estimation of Random-Coefficient Demand Models: Two Empiricists' Perspective." Review of Economics and Statistics 96.1 (2014): 34-59.
- *Nevo, Aviv. "A practitioner's guide to estimation of random-coefficients logit models of demand." Journal of economics & management strategy 9.4 (2000): 513-548.*

- Reynaert, Mathias, and Frank Verboven. "Improving the performance of random coefficients demand models: the role of optimal instruments." Journal of Econometrics 179.1 (2014): 83-98.
- Skrainka, Benjamin S., and Kenneth L. Judd. "High performance quadrature rules: How numerical integration affects a popular model of product differentiation." Available at SSRN 1870703 (2011).