BLP Turns 21

Christian Hansen

January 29, 2016

(Very) High-Level Overview of BLP

BLP - estimation of supply-demand system with aggregate data based on characteristics-based discrete choice framework

BLP addresses (at least) two concerns with earlier approaches:

(Very) High-Level Overview of BLP

BLP - estimation of supply-demand system with aggregate data based on characteristics-based discrete choice framework

BLP addresses (at least) two concerns with earlier approaches:

- Usual logit structure implies very restrictive and unrealistic substitution patterns

(Very) High-Level Overview of BLP

BLP - estimation of supply-demand system with aggregate data based on characteristics-based discrete choice framework

BLP addresses (at least) two concerns with earlier approaches:

- Usual logit structure implies very restrictive and unrealistic substitution patterns
  - Address by adding flexibility by using random coefficients (interactions between product and individual characteristics)
  - Berry and Haile “Identification in Differentiated Products Markets”

(Very) High-Level Overview of BLP

BLP - estimation of supply-demand system with aggregate data based on characteristics-based discrete choice framework

BLP addresses (at least) two concerns with earlier approaches:

- Usual logit structure implies very restrictive and unrealistic substitution patterns
  - Address by adding flexibility by using random coefficients (interactions between product and individual characteristics)
  - Berry and Haile "Identification in Differentiated Products Markets"
- Unobserved product characteristics (or other features) lead to endogeneity

(Very) High-Level Overview of BLP

BLP - estimation of supply-demand system with aggregate data based on characteristics-based discrete choice framework

BLP addresses (at least) two concerns with earlier approaches:

- Usual logit structure implies very restrictive and unrealistic substitution patterns
  - Address by adding flexibility by using random coefficients (interactions between product and individual characteristics)
  - Berry and Haile "Identification in Differentiated Products Markets"
- Unobserved product characteristics (or other features) lead to endogeneity
  - Address by using IV's and clever inversion
  - Gandhi and Houde "Measuring Substitution Patterns in Differentiated Products Industries"

BLP on product attributes

"*The choice of which attributes to include in the utility function is, of course, ad hoc.*" BLP, p. 872

BLP on instruments

"*It is important to realize that the instruments associated with product J include functions of the characteristics and cost shifters of all other products.*" BLP, p. 855

"*Though polynomial approximations are easy to compute, there is a dimensionality problem in using them to approximate functions whose arguments include the characteristics of all competing products.*" BLP, p. 860 [In reference to approximating optimal instruments]

"*Note that the dimension of the first order terms in this basis is* $3K$*, where* K *is the dimension of* z_j*. In contrast, the dimension of the first order terms in the unrestricted basis is* JK*.*" BLP, p. 861 [Following a clever dimension reduction strategy based on exchangeability. z_j are the product characteristics mentioned above. Note that K could be big and one might want more than first order terms.]

BLP on instruments

"*It is important to realize that the instruments associated with product J include functions of the characteristics and cost shifters of all other products.*" BLP, p. 855

"*Though polynomial approximations are easy to compute, there is a dimensionality problem in using them to approximate functions whose arguments include the characteristics of all competing products.*" BLP, p. 860 [In reference to approximating optimal instruments]

"*Note that the dimension of the first order terms in this basis is $3K$, where K is the dimension of z_j. In contrast, the dimension of the first order terms in the unrestricted basis is JK.*" BLP, p. 861 [Following a clever dimension reduction strategy based on exchangeability. z_j are the product characteristics mentioned above. Note that K could be big and one might want more than first order terms.]

Economic theory rarely seems to provide guidance on *exactly* which variables and functional forms one should use.

Nonparametrics

Nonparametrics (current buzzwords: "high-dimensional models," "data-mining," "big-data") tries to use researchers' beliefs AND data to address variable choice and functional form concerns.

- Aim to adapt model flexibility to data at hand (informed by beliefs)
- Want procedures/models that scale to allow for many observations and/or many predictors
  - Traditional nonparametrics (e.g. kernels, knn, series) struggle in the latter case
  - More recent approaches (e.g. Lasso methods, trees and forests) adapt more readily to high-dimensional cases (by exploiting some structure/beliefs)

Nonparametrics and Structural Estimation

Nonparametric approaches are tailored for reduced form (predictive) contexts

Need to be adapted to aid in structural estimation. E.g. (a biased list)

- linear IV: Belloni, Chen, Chernozhukov, and Hansen (2012)
- partially linear model, ATE: Belloni, Chernozhukov, and Hansen (2014), Farrell (2015)
- linear FE model: Belloni, Chernozhukov, Hansen, and Kozbur (2015)
- heterogeneous ATE: Athey and Imbens (2015), Athey and Wager (2015)
- LATE, LQTE, Moment Models: Belloni, Chernozhukov, Fernandez-Val, and Hansen (2015)
- Likelihood and moment models: Chernozhukov, Hansen, and Spindler (2015a, 2015b)
- BLP: Gillen, Moon, and Shum (2014)
- continuous random coefficient model: Hoderlein and Spindler (2015)
- nonparametric IV with shape constraints: Blundell, Horowitz and Parey (2013), Freyberger and Horowitz (2013), Chetverikov and Wilhelm (2015)

Stylized Example: Logit Demand Estimation

Model:

$$\log(s_{it}) - \log(s_{0t}) = \alpha p_{it} + x_{it}'\beta + \varepsilon_{it}$$
$$p_{it} = z_{it}'\delta + x_{it}'\gamma + u_{it}$$

- s_{it} is the market share of product i in market t. (Good 0 is the outside option)
- p_{it} is price and treated as endogenous
- x_{it} are observed included product characteristics
- z_{it} are instruments
- functional form (linearity) could be substantially relaxed - we'll be using basis expansions in a second anyway
- Goal: Learn α
- Use automobile data as in BLP in our example

Example: Baseline Results

Follow BLP just using demand side:

- 5 variables in x: constant, air conditioning, hp/weight, MP$, Size
- 10 additional instruments:
  - sum of 5 x variables for all other products from same firm
  - sum of 5 x variables for all products produced by other firms
- Estimates of α:
  - OLS: -.089 (.004) [Inelastic demand estimated for 1502 products]
  - IV: -.142 (.012) [Inelastic demand estimated for 670 products]

Why these variables and instruments?

Example: Add Some Flexibility

- Controls and instruments above plus
  - Controls: time trend, quadratics and cubics in all variables (including t), all first-order interactions ($p_x = 24$)
  - Instruments: sums of all variables as above ($p_z = 48$)
- Estimates of α
  - OLS: -.099 (.005) [Inelastic demand estimated for 1405 products]
  - IV: -.127 (.014) [Inelastic demand estimated for 874 products - a little worse than previously]

Are we overfitting? Are all of those instruments relevant? What happens when we use irrelevant instruments (many-IV and weak-IV literature)?

Example: "Adaptive" Estimation

What if we use economic intuition (and practical data limitations) to pick the set of things that might matter and then let the data help us to decide what we really need to use? [Details in Chernozhukov, Hansen, and Spindler (2015a,b)]

Case I: Data-dependent choice of controls and instruments from among baseline set ($p_x = 5$, $p_z = 10$)

- Choose to use all 5 x variables but only 3 instruments
- Estimate of α: -.208 (.015) [Inelastic demand estimated for 39 products]
- Difference probably due to many instrument bias in previous results

Example: "Adaptive" Estimation

What if we use economic intuition (and practical data limitations) to pick the set of things that might matter and then let the data help us to decide what we really need to use? [Details in Chernozhukov, Hansen, and Spindler (2015a,b)]

Case II: Data-dependent choice of controls and instruments from among augmented set ($p_x = 24$, $p_z = 48$)

- Choose to use 14 x variables but only 3 instruments (not the same as before including nonlinear term)
- Estimate of α: -.291 (.016) [Inelastic demand estimated for 1 product]
- Suggests model with only linear terms overly parsimonious
- Difference (likely) also partially accounted for by alleviating many instrument bias

More reasonable estimates of own-price elasticities using "modern" nonparametrics

Obviously not dealing with usual drawbacks of logit demand model

Potentially Interesting? Applications

- Estimating games as in, e.g., Hotz and Miller (1993), Aguirregabiria and Mira (2007), Bajari, Benkard, and Levin (2007), Pakes, Ostrovsky, and Berry (2007), Pesendorfer and Schmidt-Dengler (2008)
  - Need first step estimates of conditional distributions
- Estimating “simple” demand models in large consumer datasets (e.g. scanner data) as in, e.g., Bajari, Nekipelov, Ryan, and Yang (2015)
  - high-dimensional feature space and lots of data - want to adapt to improve demand forecast performance
  - large choice sets - do we arbitrarily focus on a set of choices and lump everything else into the outside option? Could we do better?
  - “Bundled choices”/Choices that are not purchase/don’t purchase (e.g. involve a purchase amount)

Potentially Interesting? Applications

- Learning structure over random coefficients
  - Random coefficients need not be independent normal - might want sensible, flexible regularization structure in situations where random coefficients on even moderate dimensional feature spaces are available
- (Conditional) Moment inequality models
  - Need to estimate conditional moments. Which moments are worth using? Which are just adding noise (and many moment problems)?
  - Good progress on this in Chernozhukov, Chetverikov, and Kato "Testing Many Moment Inequalities" (2014)

Parting Thoughts

Many models used in economics (including structural IO) are "high-dimensional" - We don't usually know how everything fits together up to a small set of unknown parameters

Use of sensible nonparametric methods allow one to adapt to structure in the data in a rigorous and principled way - in the spirit of BLP

- NOT talking about "bad" data-mining - unprincipled snooping for signal/patterns/significance/whatever
- NOT talking about ignoring economics, models, and structural objects in the interest of purely letting the data speak and prediction
- Do want methods with theoretical guarantees, recognize biases that result from regularization, can provide uniformly valid inference over a wide class of models (a lot of current research in econometrics/statistics is focused here)

Thank you!