Empirical Industrial Organization & Consumer Choice

2d Random Coefficient Logit Models

Summer 2018

Sebastian Kranz, Ulm University

June 28, 2018

Consumer Heterogeniety and IIA

- Recall that cross price effects in the simple logit model only depend on market shares
$$\frac{\partial s_j}{\partial p_k} = \alpha s_j s_k \forall k \neq j$$
which is a manifestation of the independence of irrelevant alternatives (IIA) property.
- Assume we observe market shares from two different types of consumers $i = 1,2$ that each make up w_i percent of the consumer population
  - e.g. consumers with income above 60000€ and consumers with income below 60000€
- Assume both consumer types have different weights β_i for product attributes (and possibly different weights α_i for prices). The representative utility of consumer type i for product j is given by
$$V_{ij} = \beta_i' x_j - \alpha_i p_j + \xi_j$$

Consumer Heterogeneity and IIA

- The market share of product j within consumer type i is

$$s_{ij} = \frac{\exp V_{ij}}{1 + \sum_{k=1}^{J} \exp V_{ik}}$$

and satisfies IIA:

$$\frac{\partial s_{ij}}{\partial p_k} = \alpha s_{ij} s_{ik} \forall k \neq j$$

- Total market shares satisfy

$$s_j = \sum_{\forall i} w_i s_{ij}$$

$$\frac{\partial s_j}{\partial p_k} = \alpha \sum_{\forall i} w_i s_{ij} s_{ik} \forall k \neq j$$

Exercise 12: Consider 4 products A,B,C, D each with total market share 25% (no outside good). We have two equally sized consumer groups $i = 1, 2$. In group 1, A and B have market shares of approximately 50% while C and D have almost zero sales. Consider a price increase of product A by 1€, by how much do the total market shares of B and C change? Show that IIA is not satisfied for total market shares and discuss the result.

Random Coefficients Logit Model

- Assumes that β_i (and possibly α_i) follow some distribution $F(\beta_i)$ whose parameters we want to estimate.
- Market shares are then given by integrating over the distribution of consumer types

$$s_j = \int \frac{\exp V_{ij}(\beta_i)}{1 + \sum_{r=1}^{J} \exp V_{ir}(\beta_i)} dF(\beta_i)$$

- This model is also called mixed logit, since market shares are a mixture of differently parametrized logit models.
- There are many ways to formulate $F(\beta_i)$ and one can possibly use known empirical distributions of certain consumer characteristics, e.g. income distribution of consumers.
- For estimation and simulation s_j is often approximated by Monte Carlo integration. E.g. sample I different types β_i from F and approximate

$$s_j = \frac{1}{I} \sum_{i=1}^{I} \frac{\exp V_{ij}(\beta_i)}{1 + \sum_{r=1}^{J} \exp V_{ir}(\beta_i)}$$

Random Coefficients Logit Model

- Most papers assume that each individual n has its own type, i.e. $i = n$
- Formulation based on Berry, Levinsohn and Pakes (1995)

$$U_{ij} = \underbrace{x_j \bar{\beta} - \alpha p_j + \xi_j}_{\delta_j} + \underbrace{\sum_k x_{jk} \sigma_k \nu_{ik}}_{\mu_{ij}} + \varepsilon_{ij}$$

  - We have
$$\beta_{ik} = \bar{\beta}_k + \sigma_k \nu_{ik}$$
  - $\alpha, \bar{\beta}$, and σ_k are the unknown parameters that we want to estimate.
  - ν_{ik} is assumed to be i.i.d. standard normally distributed, i.e. σ_k measures the variance of individual taste differences with respect for attribute x_k.

- No closed form solution for integral of market share of product j

$$\tilde{s}_j = \int \frac{\exp\left(\delta_j + \mu_{ij}(\nu)\right)}{1 + \sum_{r=1}^{J} \exp\left(\delta_r + \mu_{ir}(\nu)\right)} dF(\nu)$$

Exercise 13:

(a) Consider $K = 2$ product attributes. Use the *pseudo*-random number generator in R to randomly draw $I = 40$ consumer types from a standard normal distribution and plot them. Repeat the procedure several times to get a feeling for the variability of the random population.

(b) There alternative methods that generate so called *quasi*-random numbers that yield "more evenly allocated" samples. One famous example are so called Halton sequences. Quasi random numbers allow better approximation of the integral of market shares than pseudo-number generation. This allows more accurate estimations. Train Chapter 9.3 discusses these issues in detail. Use the R package "randtoolbox" to simulate consumer types using two dimensional Halton sequences and show the plot.

(c) Plot the resulting distribution of β_i for different values of $\bar{\beta}$ and σ in the BLP random coefficient model. IO models often distinguish between horizontal product differentiation (some consumers like product A more than B, others like B more than A) and vertical product differentiation (all consumers like product A more than B). Using your plots, discuss informally how differences in certain product attributes x_k correspond more to horizontal or vertical product differentiation and how this depends on the values of $\bar{\beta}_k$ and σ_k.

Exercise 14:

(a) Write a function that computes Bertrand equilibrium prices for the random coefficient model. Simulate some random products and consumer types and compute the equilibrium prices and market shares.

(b) Generate a plot that illustrates the position of the products in the space of attributes, their markup and their market share. Investigate and interpret informally the relationship between position, markup and market share. Investigate cross price elasticities between the products.

Random Coefficient Logit Model

- Let $\theta = (\bar{\beta}, \sigma, \alpha)$ be parameters specifying U_{ij} and the distribution of consumers i
- Let $\tilde{s}_j(\theta, \xi)$ be the market share of j as function of θ and the unobserved product demand shocks ξ
- Berry (1994) shows that for every θ there exists a unique vector of product demand shocks $\xi^*(\theta)$ such that market shares $\tilde{s}_j(\theta, \xi)$ of the model equal the vector of observed market shares s:

$$s = \tilde{s}_j(\theta, \xi^*(\theta))$$

- Recall that in simple logit model, we found

$$\log s_j - \log s_0 = x_j'\beta - \alpha p_j + \xi_j,$$

i.e.

$$\xi^*(\alpha, \beta) = \log s_j - \log s_0 - \left(x_j'\beta - \alpha p_j\right)$$

- In random coefficient model $\xi^*(\theta)$ must be computed numerically.

Least Squares Estimation

- If it would hold that
$$E(\xi | X, p) = 0$$
i.e. we had no endogeniety problem, we could consistently estimate $\theta = (\bar{\beta}, \sigma, \alpha)$ by solving the following least squares problem with a non-linear restriction:
$$\min_{\theta} \sum_{t} \sum_{j} \xi_{jt}^2$$
$$\text{subject to} s_{jt} = \tilde{s}_{jt}(\theta, \xi) \forall j, t$$
- In BLP we can approximate market shares $\tilde{s}_{jt}$ by randomly drawing a finite number I of consumer types ν_i from K-dimensional standard normal distribution and then approximate the average market shares by
$$\tilde{s}_{jt} = \frac{1}{I} \sum_{i=1}^{I} \frac{\exp\left(\delta_{jt}(\theta, \xi) + \mu_{ijt}(\theta, \nu_i)\right)}{1 + \sum_{r=1}^{J} \exp\left(\delta_{rt}(\theta, \xi) + \mu_{irt}(\theta, \nu_i)\right)}$$
- Given that endogeniety problems are likely to be present, the standard estimation method is generalized methods of moments (GMM), which yields a similar optimization problem.

Excursion: GMM Estimation

- GMM estimation is often used as a more flexible version of IV estimation (allowing non-linearities), but it can can also replicate OLS or ML estimation.
- Consider M functions $h_m(\beta)$ of observable data and coefficients β that each return T values (one for each observation in the data) that all have an expected value of 0 under the true parameters $\beta = \beta^0$

$$E(h_m(\beta^0)) = 0 \text{ for } m = 1, ..., M$$

- Typical form of h: If Z is a matrix of M valid instruments (including exogenous regressors), we get for each instrument a moment condition from

$$h_m(\beta) = Z'_m(y - X\beta)$$

since $\varepsilon = y - X\beta^0$ and $E(Z'_m\varepsilon) = 0$.

Excursion: GMM Estimation

- For estimation, one uses the sample equivalent of the the moment conditions

$$g_m(\beta) = \frac{1}{T}\sum_{t=1}^{\infty} h_m(\beta)$$

- The GMM estimator minimizes a weighted sum of the squared sample moment conditions

$$\hat{\beta}^{GMM} = \arg\min_{\beta} g'(\hat{\beta}) W g(\hat{\beta})$$

  where W is a weighting matrix.
- The ideal weighting matrix is the inverse of the covariance matrix of $h_m(\beta^0)$

$$W^{opt} = (E(h(\beta^0)h'(\beta^0)))^{-1}$$

  where expectations are taking over the realizations of y (and possibly X and Z).
  - Moment conditions that have a high variance or that are strongly correlated with other moment conditions get less weight.

Excursion: GMM Estimation

- Consider the "IV" moment conditions

$$g_m(\beta) = \frac{1}{T}\sum_{t=1}^{\infty} Z_m' \hat{\varepsilon}(\beta)$$

- If the residuals ε_t are i.i.d. distributed, it can be shown that the optimal weighting matrix is

$$W^{opt} = T\left(Z'Z\right)^{-1}$$

- The GMM estimator then solves

$$\hat{\beta}^{GMM} = \arg\min_{\beta} \frac{1}{T}\left(Z'\hat{\varepsilon}(\beta)\right)'\left(Z'Z\right)^{-1}\left(Z'\hat{\varepsilon}(\beta)\right)$$

with

$$\hat{\varepsilon}(\hat{\beta}) = y - X\hat{\beta}$$

GMM Estimation of Random Coefficient Models

- The GMM estimator in the random coefficient model is

$$\hat{\theta}^{GMM} = \arg\min_{\theta} \left(Z'\xi\right)' \left(Z'Z\right)^{-1} \left(Z'\xi\right)$$
$$\text{subject to } s_{jt} = \tilde{s}_{jt}(\theta, \xi) \forall j, t$$

- The highly non-linear constraints make this a tough optimization problem
  - Berry, Levinsohn and Pakes (1995) propose a nested fixed point algorithm. An outer loop searches candidate values for θ and an inner loop computes for each value of θ, the demand shocks $\xi^*(\theta)$ that satisfy the constraint. However, problems can arise from numerical instabilities.
  - Dube, Fox, and Su (2011) show that it can be beneficial to formulate the problem as a MPEC problem (Mathematical Programming with Economic Constraints) and solve it with commercial solvers like KNITRO, SNOPT or CONOPT.
  - Brunner, Heiss, Rohman and Weiser (2017) develop a reliable version of the nested fixed point algorithm that deals well with numerical approximation errors and implement it in the R package "BLPestimatoR"

Exercise 15: (Solution file: "code empio 2d car data.r")

(a) Estimate a BLP random coefficient logit model for our car dataset using the German market only. Use the same attributes and instruments as we used for the nested logit model. Show the estimated price elasticities for the year 1999 for a subset of cars.

Literature

- Berry, S. “Estimating Discrete-choice Models of Product Differentiation.” The RAND Journal of Economics (1994): 242–262.
- Berry, S., J. Levinsohn, and A. Pakes. “Automobile Prices in Market Equilibrium.” Econometrica: Journal of the Econometric Society (1995): 841–890.
- Nevo, A. “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand.” Journal of Economics & Management Strategy 9, no. 4 (2000): 513–548.
- Dube, J-P, Fox, T, Su, C-L “Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation”, Econometrica (2011).