ARE213 **Econometrics**

Spring 2006 UC Berkeley Department of Agricultural and Resource Economics

Discrete Response Models VI: Random-Effects Logit Models: Berry-Levinsohn-Pakes (BLP)

I. Set Up

Here we consider again random effects logit models. Such models have recently found much application in Industrial Organization, where they are used to model demand for differentiated products, often in settings with a large number of products. The first and very influential application of these methods by Berry, Levinsohn and Pakes (1995) looked at the market for automobiles.

Compared to the earlier examples we have looked at there is an emphasis in this study and those that followed it on the large number of goods and the potential endogeneity of some of the product characteristics. (Typically one of the regressors is the price of the good.) In addition the procedure only requires market level data. We do not need individual level purchase data, just market shares and estimates of the distribution of individual characteristics by market. In practice we need a fair amount of variation in these things, but in principle this is less demanding in terms of data required. On the other hand, we do need data by market, where before we just needed individual purchases in a single market (although to identify price effects we would need variation in prices by individuals in that case).

The data have three dimensions: products, indexed by $j = 0, \ldots, J$, markets, $t = 1, \ldots, T$, and individuals, $i = 1, \ldots, N_t$. We only observe one purchase per individual. The large sample approximations are based on large N and T, and fixed J.

Let us go back to the random coefficients model, now with each utility indexed by individual, product and market:

$$U_{ijt} = \beta_i' X_{jt} + \zeta_{jt} + \epsilon_{ijt}.$$

The ζ_{jt} is a unobserved product characteristic. This can include product and market dummies (for example, we can have $\zeta_{jt} = \zeta_j + \zeta_t$). Unlike the observed product characteristics this unobserved characteristic does not have a individual-specific coefficient. The observed product characteristics may include endogenous characteristics like the price.

The ϵ_{ijt} unobserved components have extreme value distributions, independent across all individuals i, products j, and markets t.

The random coefficients β_i, with dimension equal to that of the observable characteristics X_{jt}, say K, are assumed to be related to individual observable characteristics. We postulate the following linear form:

$$\beta_i = \beta + Z_i'\Gamma + \eta_i,$$

with

$$\eta_i | Z_i \sim \mathcal{N}(0, \Sigma).$$

So if the dimension of Z_i is L, then Γ is a L by K matrix. The Z_i are normalized to have mean zero, so that the β's are the average coefficients. The normality assumption is not necessary. Other distributional assumptions can be substituted.

Berry, Levinson and Pakes (19) developed an approach to estimate models of this type that does not require individual level data. Instead it exploits aggregate (market level) data in combination with estimates of the distribution of Z_i. Specifically the data consist of estimated shares $\hat{s}_{tj}$ for each choice j in each market t, combined with observations from the marginal distribution of individual characteristics (the Z_i's) for each market, often from representative data sets such as the CPS.

II. Estimation

First write the utilities as

$$U_{ijt} = \delta_{jt} + \nu_{ijt} + \epsilon_{ijt},$$

where

$$\delta_{jt} = \beta' X_{jt} + \zeta_{jt},$$

and

$$\nu_{ijt} = (Z_i'\Gamma + \eta_i)' X_{jt}.$$

Now consider for fixed Γ and Σ and δ_{jt} the implied market share for product j in market t, s_{jt}. This can be calculated analytically in simple cases. For example with $\Gamma_{jt} = 0$ and Σ an identity matrix, the market share is

$$s_{jt}(\delta_{jt}, \Gamma = 0, \Sigma = I) = \frac{\exp(\delta_{jt})}{\sum_{l=0}^{J} \exp(\delta_{lt})}.$$

More generally, this is a more complex relationship. We can always calculate the implied market share by simulation: draw from the distribution of Z_i in market t and draw from the distribution of η_i, and calculate the implied purchase probability (or even simulate the implied purchase by also drawing from the distribution of ϵ_{ijt}). Do that repeatedly and you will be able to calculate the market share for this product/market. Call the vector function obtained by stacking these functions for all products and markets $s(\delta, \Gamma, \Sigma)$.

Next, fix only Γ and Σ. For each value of δ_{jt} we can find the implied market share. Now find the vector of δ_{jt} such that the implied market shares are equal to the observed market shares $\hat{s}_{jt}$ for all j, t. BLP suggest using the following algorithm. Given a starting value for δ_{jt}^0, use the updating formula:

$$\delta^{k+1} = \delta^k + \ln s - \ln s(\delta^k, \Gamma, \Sigma).$$

BLP show this is a contraction mapping, and so it will converge to a function $\delta(s, \Gamma, \Sigma)$ expressing the δ as a function of observed market shares, and parameters Γ and Σ. Calculating this function can be quite demanding. For each iteration in the contraction mapping you need to approximate the implied market shares accurately, and then you will need to do this repeatedly to get the contraction mapping to converge.

Note that this does require that each market share is accurately estimated. If all we have is an estimated market share, then even if this is unbiased, the procedures will not necessarily work. In that case the log of the estimated share is not unbiased for the log of the true share.

Given this function $\delta(s, \Gamma, \Sigma)$ define the residuals

$$\omega_{jt} = \delta_{jt}(s, \Gamma, \Sigma) - \beta' X_{jt}.$$

At the true values of the parameters and the true market shares this is equal to the unobserved product characteristic ζ_{jt}.

Now we can use GMM or instrumental variable methods. We assume that the unobserved product characteristics are uncorrelated with observed product characteristics (other than typically price). This is not sufficient since the observed product characteristics enter directly into the model. We need more instruments, and typically use things like characteristics of other products by the same firm, or average characteristics by competing products. The general GMM machinery will also give us the standard errors for this procedure. Again, given the difficulty in evaluating $\delta(s, \Gamma, \Sigma)$, doing GMM here can be quite a challenge.

It should also be noted that this procedure is not efficient. Given the assumptions we are making here we can in principle do maximum likelihood. However, that would be even more difficult than the GMM procedure outlined above. It is not clear how much efficiency is lost doing GMM here.

It is instructive to see what this approach does if we in fact have a conditional logit model with fixed coefficients. In that case $\Gamma = 0$, and Σ is an identity matrix. Then we can invert the market share equation to get

$$\delta_{jt} = \ln s_{jt} - \ln s_{0t},$$

where we set $\delta_{0t} = 0$. (this is typically the outside good, whose utility is normalized to zero). The residual is

$$\zeta_{jt} = \delta_{jt} - \beta' X_{jt} = \ln s_{jt} - \ln s_{0t} - \beta' X_{jt}.$$

With a set of instruments W_{jt}, we run the regression

$$\ln s_{jt} - \ln s_{0t} = \beta' X_{jt} + \varepsilon_{jt},$$

using W_{jt} as instrument for X_{jt}, using as the observational unit the market share for product j in market t.

So here the technique is very transparent. It amounts to transforming the market shares to something linear in the coefficients so we can use two-stage-least-squares. More generally the transformation is going to be much more difficult with the random coefficients implying that there is no analytic solution. Computationally these things can get very complicated. Note however that we can estimate these models now without having individual level data, and that at the same time we can get a fairly flexible model for the substitution patterns. At the same time you would expect to need a lot of structure to get the parameters precisely estimated just as in the other models. Of course if you compare the current model to the nested logit model you can impose such structure by imposing restrictions on the covariance matrix.

Comparisons of the models are difficult. Obviously if the structure imposed is correct it helps, but we typically do not know what the truth is, so we cannot conclude which one is better on the basis of the data typically available.

References

Berry, S., J. Levinsohn, and A. Pakes, (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, 841-890.

Nevo, A. (2000), "A Practioner's Guide to Estimation of Random-Coefficient Logit Models of Demand," *Journal of Economics & Management Science,* Vol. 9, No. 4, 513-548.