Problem Set 3: Structural Discrete Choice Models

Chris Conlon

due: July 10

Okay this one is tricky – you are welcome to work in pairs– you can submit in pairs as long as you promise to both contribute something

Rust: The Stata Estimator (though a bit easier in R/Matlab)

This is taken from Han Hong's problem set at Stanford, the idea is that we can use the arguments in Hotz-Miller (1993), or Pesendorfer Schmidt-Dengler (2008) to construct an optimization free method to recover the utility parameters in the Rust problem.

We began by defining the choice specific value function with ϵ_{it} i.i.d. and EV.

$$
\begin{aligned}
v(x,d) &= u(x,d) + \beta \int \log\left(\sum_{d' \in D} \exp(v(x',d'))\right) p(x'|x,d)dx' \\
v(x,d) &= u(x,d) + \beta \int \log\left(\sum_{d' \in D} \exp(v(x',d') - v(x',1))\right) p(x'|x,d)dx' + \beta \int v(x,1)p(x'|x,1)dx'
\end{aligned}
$$

1. Estimate $p(x'|x,d)$ non parametrically or parametrically (for example as a set of multinomial with n outcomes or an exponential distribution). Call your estimate $\hat{p}(x'|x,d)$.
2. Estimate $p(d|x)$ (the CCP) non-parametrically. You can use the binomial logit model with a basis function (increasing number of terms) or you can use a kernel such as **ksdensity** or **ecdf**.
3. Now use the Hotz-Miller inversion to estimate: $\hat{v}(x,d) - \hat{v}(x,1) = \log \hat{p}(d|x) - \log \hat{p}(1|x)$
4. Normalize $u(x,1) = 0$ and so for $= 1$ we have that

$$
\begin{aligned}
v(x,1) &= \beta \int v(x',1)p(x'|x,1)dx' + \beta \int \log\left(\sum_{d' \in D} \exp(\hat{v}(x',d') - \hat{v}(x',1))\right) \hat{p}(x'|x,1)dx' \\
&= \beta \int v(x',1)p(x'|x,1)dx' - \beta \int \log\left(\hat{p}(1|x')\right) \hat{p}(x'|x,1)dx'
\end{aligned}
$$

This defines a fixed point that we can iterate on to obtain a nonparametric estimate of $\hat{v}(x,1)$. Add this to $\hat{v}(x,d) - \hat{v}(x,1)$ to recover the choice specific value functions for $d = 1, \ldots, D$.

5. Once we know $\hat{v}(x, d)$ for all $d \in D$ we can recover the nonparametric estimate of $u(x, d)$ for $d \geq 2$ by

$$\hat{u}(x, d) = \hat{v}(x, d) - \beta \int \log \left(\exp(\hat{v}(x', d')\right) \hat{p}(x'|x, d)dx'$$

This estimator should be very simple to implement (and only requires one fixed point) so we could do inference via the bootstrap if we wanted to.

BLP Demand Estimation

This problem set uses data on "Over the counter" Headache medicine (i.e. aspirin, tylenol and such) and was graciously provided by Vishal Singh, so if you want good demand data make friends with Marketing people![1]The data is at the store/week level for 3 brands and 3 package sizes.

- Count: Number of People that go into the store each week.
- Promotion: Is there a promotion on the product that week.
- Price: Price of the package.
- Week and Store are the time and market indicator.

Demographic data for this problem set is composed of the following:

- Income: Average income for area served by a particular store.
- This varies with stores but not across weeks.

Assume a utility specification for u_{ij}, household i's utility from brand j in store-week t:

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

where X_{jt} are characteristics of brand j, ξ_{jt} is the unobserved (to the econometrician) quality term for brand j and ε_{ijt} is a disturbance term which is identically and independently distributed with a type I extreme value.

[1]But I stole this from Alan Collard-Wexler!

Brand	Size (tab)	Market Share	Unit Price	Price/50 tab	Unit Wholesale Price
Tylenol	25	8.9%	3.52	6.91	2.26
Tylenol	50	11.1%	4.99	4.99	3.76
Tylenol	100	7.6%	7.14	3.57	5.93
Advil	25	7.3%	3.02	6.29	2.02
Advil	50	5.1%	5.19	5.19	3.70
Advil	100	2.2%	8.23	4.12	6.25
Bayer	25	2.5%	2.66	5.35	1.90
Bayer	50	2.0%	3.81	3.81	2.67
Bayer	100	4.9%	4.05	2.02	3.73
Store Brand	50	6.2%	3.57	3.57	2.48
Store Brand	100	4.2%	3.29	1.65	1.51

Table 1: Summary Statistics for Headache Data

Linear Part: Logit Estimator

Hint: It is tricky to create your own sets of multiple fixed effects in MATLAB use `dummyvar`. In R these are easier to construct. Recall that there are two ways to include FE (dummy variables and first difference).

You can report all of the following in one or two nicely organized tables.

For this setting assume that $\alpha_i = \alpha$ and $\beta_i = \beta$ (we have no random coefficients)

1. Estimate the mean utilities δ_{jt} using the inversion. Write the algebra, and store this in a variable `delta`.
2. Using OLS for δ_{jt} with price and promotion as product characteristics.
3. Same as above. Add brand FE. Write down the equation for $\Delta\xi_{jt}$ the new structural error.
4. Same as above. Add brand $\times$ store FE. Write down the equation for $\Delta\xi_{jt}$ the new structural error.
5. Estimate (2)-(4) using IV and the COST instruments only.
6. Estimate (2)-(4) using IV and the Hausman instruments only (average price in other markets).
7. Estimate (2)-(4) using IV and the both sets of instruments plus the average characteristics of other products in the same market (BLP) instruments.
8. Write the analytic formula for the Logit model and compute the mean own price elasticities by brand in the final week. Either a graph or a table. Do the estimates make sense? What is true about the "premium" brands?
9. How do the first-stage F-statistics look for each possible set of instruments?

10. Take All Quadratic interactions of all instruments. Project these interactions onto the principal components. Show how including additional principal components changes your estimates. How many principal components is enough?

Random Coefficients Logit Estimator (Extra Credit)

I provide code for this on `http://www.github.com/chrisconlon/blp-demand/`, *linked from my webpage. The code is in* ***MATLAB*** *but easy to use (hopefully).*

Now we incorporate the following two random coefficients:

- $\alpha_i = \alpha + \pi y_i$ where y_i is income. (make sure you merge this in correctly in the previous part).
- $\beta_i^B \sim N(\beta^b, \sigma_b)$ a random coefficient on a new variable `branded` which $= 1$ for branded products and $= 0$ for generic or store-brand products.

You will want to expand the set of Hausman instruments. Instead of using the average of the price of the same product in other stores during the same week, use all of the prices of the same product in other stores in the same week (or at least 30 of them).

Answer the following:

1. Do we need a random coefficients estimate a model with α_i but not β_i? Report these results for OLS and IV using all of the instruments and brand-dummies and promotion as your x variables.

2. Now suppose we also allow for a random coefficient on β_i^B the branded dummy in addition to the α_i parameter. I suggest using Gauss-Hermite quadrature rules (see `www.sparse-grids.de` for MATLAB program to generate these rules). Use brand-dummies and promotion as your x variables. You will want to estimate $[\alpha, \pi, \sigma_b]$. The code should return the standard errors.

3. Again compute the average own price elasticities (make sure to write down the correct formula which allows for random coefficients). How do the own and cross price elasticities compare to the logit model? You can show this just for a single store in a single week.

My code makes the distinction between the nonlinear parameters (which have random coefficients) which it calls ***x2*** *and the linear part of utility which it calls* ***x1****. The linear part of utility should NOT include prices, you will need a separate variable* ***price****. The instruments* ***z*** *only need to include the excluded instruments not the other regressors. The program always returns the nonlinear parameters first, followed by* ***price*** *and then the rest of* ***x1****. The Nevo Practicioner's Guide explains the linear/nonlinear parameter distinction.*