Assignment

Diego Uriarte

September 09, 2018

1 Conceptual questions on static demand and cost estimation

1.1 Explain in words why price is endogenous in Berry's framework in discrete choice models?

The price is endogenous because of the unobservable product characteristics (ξ_j). The consumer does observe these characteristics but not the econometrician (there are not in the dataset). The endogeneity occurs because the unobservable characteristics of the cars (that the consumers do observe) is obviously correlated with the price. For that reason, in the equation:

$$\ln(s_j) - \ln(s_0) = x_j \times \beta - \alpha \times p_j + \xi_j$$

A higher ξ_j will cause a higher p_j (because firms are in a Bertrand price competition), all else equal.

1.2 Explain in words why within-group share is endogenous in discrete choice models?

The within-group is endogenous because is correlated with the unobservable product characteristics (ξ_j). For instance, if the share of Mercedes Benz automobiles goes down (maybe because there is an increase in their prices), we would expect that the share of BMW cars goes up by much more than the share of Nissan cars. This would be because both Mercedes and BMW share characteristics that appeal to the same costumers.

1.3 Derive marginal costs for single-product firms for Berry's nested logit model, given data and estimated parameters.

First, profit for firm j are given by:

$$\pi_j(\mathbf{p}, z, \xi, \omega_j, \theta) = p_j M s_j(x, \xi, p, \theta_d) - C_j(q_j, w_j, \omega_j, \gamma)$$

Where, p are prices, x, w observed characteristics, ξ and ω unobserved characteristics, θ_d demand parameter, M is total market size.

Since firms are price setters, and if we assume that there is an interior equilibrium, the FOC are:

$$p_j = c_j + s_j / |\partial s_j / \partial p_j|$$

Also, using the chain rule ($\frac{\partial s_j}{\partial \delta_j}\frac{\partial \delta_j}{\partial p_j} = \frac{\partial s_j}{\partial p_j}$), and using the definition of δ_j to differentiate $\frac{\partial \delta_j}{\partial p_j} = -\alpha$ we get:

$$-\alpha \frac{\partial s_j}{\partial \delta_j} = \frac{\partial s_j}{\partial p_j}$$

Also, we assume that the marginal cost is linear in the unobservable cost term ω_j, we get $c_j = \bar{c}(q_j, w_j, \gamma) + \omega_j$. Replacing in the FOC:

$$p_j = \bar{c_j} + \frac{1}{\alpha}[s_j/|\partial s_j/\partial \delta_j|] + \omega_j$$

Using the market share equation for the logit model, we get:

$$\partial \int_j/\partial \delta_j = \frac{1}{1-\sigma} s_j[1 - \sigma s_{\bar{j}/g} - (1-\sigma)s_j]$$

Finally, replacing in the previous expression:

$$p_j = \bar{c_j} + \frac{1-\sigma}{\alpha}/[1 - \sigma s_{\bar{j}/g} - (1-\sigma)s_j] + \omega_j$$

Since the parameters have been estimated, we can determine average marginal cost from the previous expression. These derivation was taken from Berry (1994).

1.4 What are two useful source of instruments for discrete choice models? Explain.

One possibility is to use price variation across cities or regions as cost shifters. These are a valid instrument if the marginal costs are correlated across cities, but not the unobservable product characteristics (ξ_j). That is to say that the price of a good in city j is correlated with the marginal cost in city j, that is correlated with the marginal cost in city i. But ξ_i and ξ_j are not correlated.

For the automobile sector, changes in fuel prices generate cost variation. The consumer is interested in how long can she travel with 1$ of fuel (miles per dollar), but for the car maker, increase the miles per gallon for a given model is costly. For that reason, fuel prices variation can be used as cost shifter, however they may not be that strong.

1.5 What are two reasons to estimate the supply side models along with the demand side of discrete choice models? Explain.

Only estimating the demand side is consistent with the endogeneity of prices and with equilibrium results, however, by adding the supply side we can get better or more precise estimates of the parameters. Another reason to incorporate supply side models and solve for the equilibrium is that we can test with type of competitive model fits best the data. As Miller and Weinberg (2017) used this approach to reject the hypothesis that there was a Nash-Bertrand equilibrium between two brewing company after their merge.

1.6 What are the advantages of using a random coefficients model, à la BLP, instead of a nested logit model?

IV estimator for the nested logit model is just a special case of the BLP model, where ξ_j is linear in the parameters. The advantage of BLP is that it allows for non-linear relation between ξ_j and the parameters, by recovering the function numerically. Another advantage is that BLP added equilibrium behavior to the models, and as it was discussed in the previous question, it can improve the accuracy of the parameters and allows us to test for market behavior.

2 Computational questions on static demand and cost estimation (40 points)

2.1 Collapse the data to the product / market level and drop the outside good for each market as observations, creating an "outside good share" variable for each inside good choice. Report summary statistics on market shares and outside good share by plan number. Report the number of observations.

First, we explore the dataset

```r
head(data)
```

```
## # A tibble: 6 x 11
##   individual market  plan num_plans choice price    xsi x_constant
##        <int>  <int> <int>     <int>  <int> <dbl>  <dbl>      <int>
## 1          1      1     0         5      0 0     0               0
## 2          1      1     1         5      1 1.09  0.631           1
## 3          1      1     2         5      0 1.32  0.921           1
## 4          1      1     3         5      0 0.876 0.0309          1
## 5          1      1     4         5      0 0.908 0.0125          1
## 6          2      1     0         5      0 0     0               0
## # ... with 3 more variables: x_coinsurance <dbl>, x_deductible <dbl>,
## #   x_oopmax <dbl>
```

There are 89801 persons in the dataset, 100 markets, each with different plans. We are given information about the plan chosen by each individual (`choice` variable), as well as the price, and the observable characteristic in each plan (`x_constant, x_coinsurance, x_deductible, x_oopmax`). It is important to note that plan x in market i is not the same as plan x in market j. The value `xsi` represents the unobservable characteristics, and as such, they will only be used at the end as benchmark for our regression results.

We are asked to collapse the data to the market level.The first 10 values of the collapsed dataset are showed. The complete dataset transformation can be found in the replication file.

```
## # A tibble: 10 x 11
## # Groups:   market [3]
##     market  plan   s_0    s_j s_j_given_g price     xsi x_constant
##      <int> <int> <dbl>  <dbl>       <dbl> <dbl>   <dbl>      <dbl>
##  1       1     1 0.352 0.183       0.283  1.09   0.631           1
##  2       1     2 0.352 0.300       0.463  1.32   0.921           1
##  3       1     3 0.352 0.0679      0.105  0.876  0.0309          1
##  4       1     4 0.352 0.0970      0.150  0.908  0.0125          1
##  5       2     1 0.349 0.101       0.155  0.932  0.237           1
##  6       2     2 0.349 0.0166      0.0256 0.814 -0.650           1
##  7       2     3 0.349 0.264       0.406  1.21   0.681           1
##  8       2     4 0.349 0.256       0.393  1.23   0.657           1
##  9       2     5 0.349 0.0135      0.0208 0.813 -0.738           1
## 10       3     1 0.549 0.242       0.536  1.34   0.0325          1
## # ... with 3 more variables: x_coinsurance <dbl>, x_deductible <dbl>,
## #   x_oopmax <dbl>
```

`s_0` is the outside share for each market, `s_j` is the share within market for a given plan, `s_j_given_g` is the within market share.

Now, we find the summary statistic by plan number:

	plan: 1 (N = 100)	plan: 2 (N = 100)	plan: 3 (N = 66)	plan: 4 (N = 39)	plan: 5 (N = 20)
price					
min	0.812888	0.814086	0.805460	0.813620	0.807153
max	2.13393	2.47220	1.83560	1.29913	1.47126
mean (sd)	1.19 ± 0.29	1.19 ± 0.33	1.08 ± 0.25	0.96 ± 0.13	0.95 ± 0.17
xsi					
min	-1.059690	-1.048210	-1.054930	-1.114970	-0.954659
max	1.394570	1.413610	1.462110	0.708806	0.688833
mean (sd)	0.05 ± 0.41	0.02 ± 0.51	-0.02 ± 0.52	-0.12 ± 0.42	-0.13 ± 0.47
x__coinsurance					
min	0.812888	0.814086	0.805460	0.813620	0.807153
max	2.13393	2.47220	1.83560	1.29913	1.47126
mean (sd)	1.19 ± 0.29	1.19 ± 0.33	1.08 ± 0.25	0.96 ± 0.13	0.95 ± 0.17
x__deductible					
min	0.812888	0.814086	0.805460	0.813620	0.807153
max	2.13393	2.47220	1.83560	1.29913	1.47126
mean (sd)	1.19 ± 0.29	1.19 ± 0.33	1.08 ± 0.25	0.96 ± 0.13	0.95 ± 0.17
x__oopmax					
min	0.812888	0.814086	0.805460	0.813620	0.807153
max	2.13393	2.47220	1.83560	1.29913	1.47126
mean (sd)	1.19 ± 0.29	1.19 ± 0.33	1.08 ± 0.25	0.96 ± 0.13	0.95 ± 0.17

2.2 Construct as instruments the within-group sum of every characteristic. Report summary statistics on your instruments, with the Stata "summarize" command or analog in the software that you use

First, we add a column to the table with the within-group sum of every characteristic (`z1, z2, z3`). However, following Berry (1994), we will also try with the within-group sum of the characteristics without the plan `j`. We call these three instruments (`z11, z22, z33`). Here, we report the summary statistics

Summary Statistics	N = 325
z1	
min	0.0686351
median (IQR)	0.93 (0.67, 1.23)
mean (sd)	0.95 ± 0.36
max	1.562735
z2	
min	0.2497498
median (IQR)	0.95 (0.67, 1.24)
mean (sd)	0.99 ± 0.41
max	1.935549
z3	
min	0.09085784
median (IQR)	0.89 (0.63, 1.30)
mean (sd)	0.92 ± 0.39
max	1.75793
z11	
min	0.00251275
median (IQR)	0.68 (0.42, 0.98)
mean (sd)	0.69 ± 0.35
max	1.507379
z22	
min	0.00376132
median (IQR)	0.69 (0.42, 0.99)
mean (sd)	0.72 ± 0.39
max	1.754211
z33	
min	-0.2394983
median (IQR)	0.64 (0.36, 0.96)
mean (sd)	0.65 ± 0.42
max	1.602361

2.3 Estimate a nested logit model using Berry's method, not instrumenting for within-group share or price. Report your results.

Now, we estimate the nested logit model. The model is as follows:

$$\ln(s_j) - \ln(s_0) = \beta_0 + \beta_1 x_{coinsurance} + \beta_2 x_{deductible} + \beta_3 x_{oopmax} - \alpha p_j + \sigma \ln(s_{j/g}) + \xi_j$$

I perform the first regression, using the transformed data.

```
##
## Call:
## lm(formula = ln_sj_minus_s0 ~ x_coinsurance + x_deductible +
##     x_oopmax + price + ln_sj_g, data = transformed_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98191 -0.25049  0.02532  0.24990  0.81439
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)   -0.11916    0.17393  -0.685   0.4938
## x_coinsurance -0.26490    0.14029  -1.888   0.0599 .
## x_deductible  -0.20614    0.13369  -1.542   0.1241
## x_oopmax       0.20636    0.13156   1.569   0.1177
## price          0.10443    0.10094   1.035   0.3016
## ln_sj_g        0.80387    0.03125  25.727   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3388 on 319 degrees of freedom
## Multiple R-squared:  0.8435, Adjusted R-squared:  0.8411
## F-statistic:   344 on 5 and 319 DF,  p-value: < 2.2e-16
```

2.4 Estimate a nested logit model using Berry's method, instrumenting for within-group share but not price. Report your results.

Now, I utilize 2SLS with three instrumental variables (`z1, z2, z3`) and withing-group share as endogenous.

```
##
## Call:
## ivreg(formula = ln_sj_minus_s0 ~ x_coinsurance + x_deductible +
##     x_oopmax + price + ln_sj_g | . - ln_sj_g + z1 + z2 + z3,
##     data = transformed_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0257 -0.2556  0.1309  0.4060  1.1356
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.74824    0.87511  -4.283 2.44e-05 ***
## x_coinsurance -0.02080    0.25583  -0.081   0.9352
## x_deductible  -0.28211    0.23870  -1.182   0.2381
## x_oopmax       0.62764    0.25283   2.482   0.0136 *
## price          2.11343    0.48744   4.336 1.95e-05 ***
## ln_sj_g       -0.01853    0.19364  -0.096   0.9238
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6033 on 319 degrees of freedom
## Multiple R-Squared: 0.5038,  Adjusted R-squared: 0.496
## Wald test: 66.71 on 5 and 319 DF,  p-value: < 2.2e-16
```

2.5 Estimate a nested logit model using Berry's method, instrumenting for within-group share and price. Report your results. Does it appear that price was endogenous? How are you making this judgment?

Finally, the same as in the previous question but instrumenting also by price.

```
##
## Call:
```

```
## ivreg(formula = ln_sj_minus_s0 ~ x_coinsurance + x_deductible +
##     x_oopmax + price + ln_sj_g | . - price - ln_sj_g + z1 + z2 +
##     z3, data = transformed_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11665 -0.35187 -0.08585  0.26870  2.65404
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.54694    1.83910   1.385  0.16706
## x_coinsurance -0.95821    0.34676  -2.763  0.00605 **
## x_deductible  -0.49901    0.23739  -2.102  0.03633 *
## x_oopmax       0.06652    0.28436   0.234  0.81518
## price         -1.83602    1.12769  -1.628  0.10449
## ln_sj_g        0.93102    0.30934   3.010  0.00282 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5829 on 319 degrees of freedom
## Multiple R-Squared: 0.5368,  Adjusted R-squared: 0.5295
## Wald test: 12.43 on 5 and 319 DF,  p-value: 4.953e-11
```

I observe that when I don't use instruments for price nor withing group share, the price coefficient is positive (0.1) which would indicate that an increase in price increases the market share of the plan. Controlling for within group share does not disappear this effect since the price coefficient for the second regression is 2.11. However, controlling for price does make the price coefficient negative, which intuitively makes sense since from the model we are quite sure that price is correlated with ξ_j.

In addition, using as instruments for each plan `k` the within sum characteristic for values different than `k` ($z_{11} = \sum_{i \neq k} x_i$), provides the same estimates:

```
##
## Call:
## ivreg(formula = ln_sj_minus_s0 ~ x_coinsurance + x_deductible +
##     x_oopmax + price + ln_sj_g | . - price - ln_sj_g + z11 +
##     z22 + z33, data = transformed_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11665 -0.35187 -0.08585  0.26870  2.65404
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.54694    1.83910   1.385  0.16706
## x_coinsurance -0.95821    0.34676  -2.763  0.00605 **
## x_deductible  -0.49901    0.23739  -2.102  0.03633 *
## x_oopmax       0.06652    0.28436   0.234  0.81518
## price         -1.83602    1.12769  -1.628  0.10449
## ln_sj_g        0.93102    0.30934   3.010  0.00282 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5829 on 319 degrees of freedom
```

```
## Multiple R-Squared: 0.5368,  Adjusted R-squared: 0.5295
## Wald test: 12.43 on 5 and 319 DF,  p-value: 4.953e-11
```

3 Conceptual questions on applications of discrete choice models to antitrust (20 points)

3.1 How did Prof. Nevo argue that the nested logit model was a useful demand model, in the Aetna/Humana merger case? (Please read the judge's decision in this case that is on the syllabus.)

Prof. Nevo used the CMS data on Medicare Advantage plan enrollments that also included seniors who chose Original Medicare Options. The nested logit was useful in this context because it allows us to test *"whether, an to what degree, a senior might prefer"a Medicare Advantage plan because it is a Medicare Advantage plan"* (United States District Court for the District of Columbia (2017)). The key parameter in the model is the nesting parameter, that indicates the strength of this preference and it can have values between 0 and 1.

Nevo found that 70% of users of one of the Medicare Advantage plans would change to another Medicare Advantage plan (nesting parameter of 0.65). Nevo considered this to be an conservative estimated since the data showed that around 80% of seniors change a Medicare Advantage Plan of another Medicare Advantage Plan. Using this model, he was able to prove that an hypothetical profit-maximizing monopolist could profitably increase prices in any county.

One turning point for the judge to accept Nevo's econometric evidence was that he used the defendant nesting parameters (that were much lower that his) and the lowest parameters he could find the literature, and still found that the Medicare Advantage passed the SSNIP test of 5% or 10%, that is, an hypothetical monopolist in that market could profitably increase prices in more than 5%.

3.2 How would you would perform a hypothetical monopolist test for your estimated model of insurance demand? Explain with words and equations.

I would calculate the own-price elasticity using the parameters from the model:

$$\eta_{jj} = \alpha_j p_j [s_j - \frac{1}{1-\sigma} + \frac{\sigma}{1-\sigma} s_{j/g}]$$

for each product in each market. Then, using the Hypothetical Monopolist test, I would apply a 5% increase in prices for plan j, that would result in an decrease of $5\% \times \eta_{jj}$ in quantity demanded. Then, we could check if the increase in prices is still profitable, using the marginal costs for each plan.

References

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics*. JSTOR, 242–62.

Miller, Nathan H, and Matthew C Weinberg. 2017. "Understanding the Price Effects of the Millercoors Joint Venture." *Econometrica* 85 (6). Wiley Online Library: 1763–91.

United States District Court for the District of Columbia. 2017. "United States et al. v. AETNA INC et al." https://www.justice.gov/opa/press-release/file/930361/download.