Assignment

Diego Uriarte

31 de agosto de 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.

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

One posibility is to use price variation across cities or regions as cost shifters. They are a valid instrument if the marginal costs are correlated across cities, but not the unobservable product charactheristics (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 fue (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, but 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 demande side is consistent with the endogenoity of prices and with equilibirum 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 aprooach 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 advantange 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 behaviour.

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

```
data <- as_tibble(data)
str(data)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    381768 obs. of  11 variables:
##  $ individual   : int  1 1 1 1 1 2 2 2 2 2 ...
##  $ market       : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ plan         : int  0 1 2 3 4 0 1 2 3 4 ...
##  $ num_plans    : int  5 5 5 5 5 5 5 5 5 5 ...
##  $ choice       : int  0 1 0 0 0 0 0 1 0 0 ...
##  $ price        : num  0 1.094 1.319 0.876 0.908 ...
##  $ xsi          : num  0 0.6312 0.9206 0.0309 0.0125 ...
##  $ x_constant   : int  0 1 1 1 1 0 1 1 1 1 ...
##  $ x_coinsurance: num  0 0.4596 0.4411 0.4561 0.0803 ...
##  $ x_deductible : num  0 0.1749 0.3555 0.0952 0.1625 ...
##  $ x_oopmax     : num  0 0.0973 0.3058 0.089 0.4978 ...
```

```
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 differents plans We are given information about the plan choosed by each individual (`choice` variable), as well as the price, and the observable charactheristic 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 charactheristics, and as such, they will only be used at the ended as a benchamark for our regression results.

We are asked to collapse the data to the market level. First, we will keep only the row corresponding to the choice made by the individual.

We create a table with the outside share for each market (there should be 100 values)

```r
freq_market <- data %>% filter(choice == 1) %>%
  group_by(market, plan) %>%
  summarise(n = n()) %>%
  mutate(freq = n / sum(n))

outside_market <- freq_market %>% filter(plan == 0)
```

We transform the original data, dropping the outside good and only keeping the product choose by each individual in given market:

```r
data %>% filter(choice == 1) %>%
  left_join(outside_market %>% select(market, freq), by = "market")
```

```
## # A tibble: 89,801 x 12
##    individual market  plan num_plans choice price   xsi x_constant
##         <int>  <int> <int>     <int>  <int> <dbl> <dbl>      <int>
##  1          1      1     1         5      1  1.09 0.631          1
##  2          2      1     2         5      1  1.32 0.921          1
##  3          3      1     2         5      1  1.32 0.921          1
##  4          4      1     2         5      1  1.32 0.921          1
##  5          5      1     0         5      1  0     0            0
##  6          6      1     2         5      1  1.32 0.921          1
##  7          7      1     2         5      1  1.32 0.921          1
##  8          8      1     2         5      1  1.32 0.921          1
##  9          9      1     2         5      1  1.32 0.921          1
## 10         10      1     2         5      1  1.32 0.921          1
## # ... with 89,791 more rows, and 4 more variables: x_coinsurance <dbl>,
## #   x_deductible <dbl>, x_oopmax <dbl>, freq <dbl>
```

Since there is information not needed, we drop them from the table and the change the `freq` variable to `s_0` (outside option):

```r
data_with_s0 <- data %>% filter(choice == 1) %>%
  left_join(outside_market %>% select(market, freq), by = "market") %>%
  select(individual:plan, s_0 = freq, price:x_oopmax)

data_with_s0
```

```
## # A tibble: 89,801 x 10
##    individual market  plan   s_0 price   xsi x_constant x_coinsurance
##         <int>  <int> <int> <dbl> <dbl> <dbl>      <int>         <dbl>
##  1          1      1     1 0.352  1.09 0.631          1         0.460
##  2          2      1     2 0.352  1.32 0.921          1         0.441
##  3          3      1     2 0.352  1.32 0.921          1         0.441
##  4          4      1     2 0.352  1.32 0.921          1         0.441
##  5          5      1     0 0.352  0     0             0         0
##  6          6      1     2 0.352  1.32 0.921          1         0.441
##  7          7      1     2 0.352  1.32 0.921          1         0.441
##  8          8      1     2 0.352  1.32 0.921          1         0.441
##  9          9      1     2 0.352  1.32 0.921          1         0.441
## 10         10      1     2 0.352  1.32 0.921          1         0.441
## # ... with 89,791 more rows, and 2 more variables: x_deductible <dbl>,
## #   x_oopmax <dbl>
```

Then, we collapse the data by calculating the shares of each product for each market.

```
data %>% filter(choice == 1) %>%
  group_by(market, plan) %>%
  summarise(n = n()) %>%
  mutate(freq = n / sum(n)) %>% filter(plan != 0)
```

```
## # A tibble: 325 x 4
## # Groups:   market [100]
##    market  plan     n   freq
##     <int> <int> <int>  <dbl>
##  1      1     1   170 0.183
##  2      1     2   278 0.300
##  3      1     3    63 0.0679
##  4      1     4    90 0.0970
##  5      2     1    97 0.101
##  6      2     2    16 0.0166
##  7      2     3   254 0.264
##  8      2     4   246 0.256
##  9      2     5    13 0.0135
## 10      3     1   214 0.242
## # ... with 315 more rows
```

```
collapsed_data <- data_with_s0  %>% group_by(market, plan) %>%
  filter(plan != 0) %>%
  summarise(s_0 = mean(s_0),
            price = mean(price),
            xsi = mean(xsi),
            x_constant = mean(x_constant),
            x_coinsurance = mean(x_coinsurance),
            x_deductible = mean(x_deductible),
            x_oopmax = mean(x_oopmax)) %>%
  left_join(freq_market %>% select(market, plan, s_j = freq), by = c("market","plan")) %>%
  mutate(s_j_given_g = s_j / (1 - s_0)) %>%
  select(market:s_0, s_j, s_j_given_g, price:x_oopmax)

collapsed_data
```

```
## # A tibble: 325 x 11
## # Groups:   market [100]
##    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 315 more rows, and 3 more variables: x_coinsurance <dbl>,
## #   x_deductible <dbl>, x_oopmax <dbl>
```

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 Barry, we will also try with the within-group sum of the characteristics without the plan `j`. We call these three instruments (`z11. z22. z33`).

```
data_with_instruments <- collapsed_data %>% group_by(market) %>%
  summarise(z1 = sum(x_coinsurance), z2 = sum(x_deductible), z3 = sum(x_oopmax)) %>%
  right_join(collapsed_data, by = "market") %>%
  select(market, plan:x_oopmax, z1:z3) %>%
  mutate(z11 = z1 - x_coinsurance, z22 = z2 - x_deductible, z33 = z3 - x_deductible)

data_with_instruments %>% select(z1, z2, z3, z11, z22, z33) %>% summary
```

```
##        z1                 z2               z3                z11
##  Min.   :0.06864   Min.   :0.2497   Min.   :0.09086   Min.   :0.002513
##  1st Qu.:0.66646   1st Qu.:0.6735   1st Qu.:0.63388   1st Qu.:0.415790
##  Median :0.92851   Median :0.9460   Median :0.89035   Median :0.678881
##  Mean   :0.94628   Mean   :0.9877   Mean   :0.91782   Mean   :0.685327
##  3rd Qu.:1.23344   3rd Qu.:1.2356   3rd Qu.:1.30162   3rd Qu.:0.976185
##  Max.   :1.56273   Max.   :1.9355   Max.   :1.75793   Max.   :1.507379
##       z22                z33
##  Min.   :0.003761   Min.   :-0.2395
##  1st Qu.:0.417792   1st Qu.: 0.3585
##  Median :0.692470   Median : 0.6365
##  Mean   :0.716415   Mean   : 0.6465
```

```
##   3rd Qu.:0.993354   3rd Qu.: 0.9571
##   Max.   :1.754211   Max.   : 1.6024
```

```
our_summary2 <-
  with(data_with_instruments,
       list("z1" = tab_summary(z1),
            "z2" = tab_summary(z2),
            "z3" = tab_summary(z3))
  )


whole_table <- summary_table(data_with_instruments, our_summary2)
```

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

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

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, we 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 parameter 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%.

References

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.