ECO310 - Tutorial 7

Estimating Marginal Costs

Francis Guiton

March 15, 2021

In this week's tutorial, we will be focusing on how to estimate marginal costs using the Verboven automobile dataset. We will consider both the standard logit model and the nested logit model, and we will assume Bertrand competition in order to pin down the price-cost margin and the marginal cost. Next week, we will see how to implement a simple counterfactual merger analysis using the "mergersim" command in STATA (type *search mergersim* in the Command Window in order to find and install the command).

1 Set-Up

We load the Verboven dataset onto STATA and generate the logarithms of our variables of interest: price (in euros), quantity, population and GDP.

```
. // Generating log-variables
. gen ln_p = ln(eurpr)

. gen ln_q = ln(qu)

. gen ln_pop = ln(pop)

. gen ln_gdp = ln(ngdp)
```

Since we will be working with the logit model, we must generate our market shares. As we have done in the previous tutorial, we generate a market share for each car model in a given market and year, and create a variable that sums these market shares across markets and years. This will then allow us to generate the outside shares. In this tutorial we will consider demand at the household level (i.e. market size = population/4). Finally, given these market shares, we generate our dependent variable - the log-odds ratio.

```
. // Creating market shares
. gen market_size = pop/4

. gen share = qu/market_size

. egen sum_share = sum(share), by(ma ye)

. gen outside_share = 1-sum_share

.
. // Generating the log-odds ratio (our dependent variable)
. gen sj_s0 = ln(share/outside_share)
```

2 Demand Estimation

Nested Logit

In order to implement the nested logit model, we must first decide how to separate the products into groups (or "nests"). In this tutorial, we will group the automobiles into two groups - foreign cars and domestic cars, given by the variable "home" (1 if domestic, 0 otherwise). In a given market and year, we generate a variable "nest_sum" that sums the market shares of each nest. Given this new variable, we compute the (log-) within-group share by dividing the individual market shares by the "nest_sum" variable.

```
. // First, we generate the sum of market shares of each nest given a market and a year
. egen nest_sum = sum(share), by(ma ye home)

.
. // We can now generate the (log) within-group market share
. gen within_share = ln(share/nest_sum)
```

We now run the nested logit model using the "reghdfe" command from last week's tutorial in order to control for multiple levels of fixed effects in our panel. Unlike last week's tutorial, here we will control for market, year, and brand fixed effects (command option "a(ma ye brd)"). We include the log- within-group share as an explanatory variable, population and GDP as covariates, and product attributes. As usual, we include the "vce(robust)" option in our command.

```
. reghdfe sj_s0 ln_p within_share hp li wi cy le he ln_pop ln_gdp, vce(robust) a(ma ye brd)
(MWFE estimator converged in 7 iterations)

HDFE Linear regression                           Number of obs     =     11,549
Absorbing 3 HDFE groups                          F(  10,   11466) =   17778.91
                                                 Prob > F          =     0.0000
                                                 R-squared         =     0.9524
                                                 Adj R-squared     =     0.9521
                                                 Within R-sq.      =     0.9380
                                                 Root MSE          =     0.3287

------------------------------------------------------------------------------
             |               Robust
       sj_s0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_p |   .0347379    .028136     1.23   0.217    -.0204135    .0898893
within_share |   1.059716   .0030231   350.54   0.000      1.05379    1.065641
          hp |   .0022768   .0004699     4.85   0.000     .0013557    .0031979
          li |   .0005364   .0042338     0.13   0.899    -.0077626    .0088354
          wi |  -.0042111   .0009483    -4.44   0.000    -.0060699   -.0023523
          cy |   7.53e-06   .0000222     0.34   0.734    -.0000359     .000051
          le |   .0002348   .0002179     1.08   0.281    -.0001922    .0006619
          he |   .0009742   .0009389     1.04   0.299    -.0008661    .0028145
      ln_pop |   -.162994   .0666984    -2.44   0.015    -.2937343   -.0322537
      ln_gdp |   .4837022   .0182389    26.52   0.000     .4479509    .5194535
       _cons |  -13.66916   1.364237   -10.02   0.000    -16.34329   -10.99502
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Categories  - Redundant  = Num. Coefs |
-------------+---------------------------------------|
          ma |         5           0           5     |
          ye |        30           1          29     |
         brd |        40           1          39    ?|
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

The ”within_share” coefficient measures the consumers’ preference correlation for products within nests. Since this estimate is greater than 1, our nested logit model is not valid. We could consider separating products into new groups and subgroups.

Standard Logit

Since our previous nested logit model is not valid, we will instead consider the standard logit model of last week's tutorial.

```
. reghdfe sj_s0 ln_p hp li wi cy le he ln_pop ln_gdp, vce(robust) a(ma ye brd)
(MWFE estimator converged in 7 iterations)

HDFE Linear regression                            Number of obs   =     11,549
Absorbing 3 HDFE groups                           F(   9,  11467) =     341.86
                                                  Prob > F        =     0.0000
                                                  R-squared       =     0.4018
                                                  Adj R-squared   =     0.3975
                                                  Within R-sq.    =     0.2203
                                                  Root MSE        =     1.1656

------------------------------------------------------------------------------
             |               Robust
       sj_s0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_p |  -1.164701   .1038432   -11.22   0.000    -1.368251   -.9611507
          hp |  -.0137051   .0017671    -7.76   0.000    -.0171689   -.0102413
          li |  -.0415553   .0136634    -3.04   0.002    -.0683379   -.0147726
          wi |   .0638831   .0033199    19.24   0.000     .0573755    .0703907
          cy |  -.0006896    .000084    -8.21   0.000    -.0008542    -.000525
          le |  -.0000936   .0007787    -0.12   0.904      -.00162    .0014327
          he |   -.017626   .0030341    -5.81   0.000    -.0235734   -.0116785
      ln_pop |    .349325   .2300971     1.52   0.129    -.1017048    .8003547
      ln_gdp |   .2826532    .062891     4.49   0.000     .1593761    .4059302
       _cons |  -17.20771   4.726521    -3.64   0.000     -26.4725   -7.942923
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Categories  - Redundant  = Num. Coefs |
-------------+---------------------------------------|
          ma |         5           0           5     |
          ye |        30           1          29     |
         brd |        40           1          39    ?|
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

3 Estimating Marginal Cost

Price-Cost Margin

Using the parameter estimate above for the (log-) price, we will compute the price-cost margin under Bertrand competition, Here we will consider the "standard" functional form for the marginal cost of firms outlined in the lecture slides. Note that given our model, price enters the consumer's utility function as a logarithm. Therefore, with $\alpha < 0$, we have:

$$\frac{\partial \sigma_i}{\partial p_i} = \frac{\alpha}{p_i} s_i (1 - s_i)$$

The above entails that our price-cost margin can be written as

$$-\frac{s_i}{\partial \sigma_i / \partial p_i} = -\frac{p_i}{\alpha(1 - s_i)}$$

We therefore compute the price-cost margin in STATA in the following way:

```
. gen price_cost_margin = -eurpr/(_b[ln_p]*(1-share))
```

Marginal Cost

Given our price-cost margin, we can now estimate the firms' marginal costs. First, since $MR_i = p_i + \frac{s_i}{\partial\sigma_i/\partial p_i}$, we compute (log-) marginal revenue as:

```
. // Using the price-cost margin, we can compute marginal revenue (in euros)
. gen mr = eurpr - price_cost_margin

. gen ln_mr = ln(mr)
```

Then, given our functional form for marginal cost, we run a regression of (log-) marginal revenue on (log-) quantity and product characteristics in order to obtain an estimate for marginal cost using the regression's fitted value.

```
. reghdfe ln_mr ln_q hp li wi cy le he ln_pop ln_gdp, vce(robust) a(ma ye brd) res
(MWFE estimator converged in 7 iterations)

HDFE Linear regression                            Number of obs   =     11,549
Absorbing 3 HDFE groups                           F(   9,  11467) =    8552.44
                                                  Prob > F        =     0.0000
                                                  R-squared       =     0.9724
                                                  Adj R-squared   =     0.9723
                                                  Within R-sq.    =     0.8860
                                                  Root MSE        =     0.1100

------------------------------------------------------------------------------
             |               Robust
       ln_mr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ln_q |  -.0188848   .0009599   -19.67   0.000    -.0207663   -.0170032
          hp |   .0084214   .0001798    46.84   0.000      .008069    .0087738
          li |   .0001027   .0013618     0.08   0.940    -.0025666     .002772
          wi |   .0034008     .00034    10.00   0.000     .0027343    .0040674
          cy |   .0000968   8.75e-06    11.06   0.000     .0000796    .0001139
          le |   .0022009   .0000757    29.08   0.000     .0020526    .0023493
          he |  -.0033937   .0002729   -12.44   0.000    -.0039286   -.0028587
      ln_pop |   .4230247   .0200337    21.12   0.000     .3837551    .4622942
      ln_gdp |   .1295281   .0050102    25.85   0.000     .1197073     .139349
       _cons |  -5.786087   .4045117   -14.30   0.000    -6.578999   -4.993175
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Categories  - Redundant  = Num. Coefs |
-------------+---------------------------------------|
          ma |         5           0           5     |
          ye |        30           1          29     |
         brd |        40           1          39    ?|
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

Unfortunately, the regression suffers from an endogeneity problem, as our log-quantity is likely correlated with the unobservables. One option would be to use an instrumental variables approach in order to control for this endogeneity. In this tutorial, we will instead consider a constant marginal cost specification, which will get rid of the potential endogeneity. Below, we regress log- marginal revenue on product characteristics. We add the "res" option to our "reghdfe" command in order to tell STATA that we will be using the parameter estimates for post-estimation commands.

```
. reghdfe ln_mr hp li wi cy le he ln_pop ln_gdp, vce(robust) a(ma ye brd) res
(MWFE estimator converged in 7 iterations)

HDFE Linear regression                            Number of obs   =     11,549
Absorbing 3 HDFE groups                           F(   8,  11468) =    8855.84
                                                  Prob > F        =     0.0000
                                                  R-squared       =     0.9713
                                                  Adj R-squared   =     0.9711
                                                  Within R-sq.    =     0.8814
                                                  Root MSE        =     0.1122

------------------------------------------------------------------------------
             |               Robust
       ln_mr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          hp |   .0088716   .0001872    47.38   0.000     .0085045    .0092386
          li |   .0008985   .0014203     0.63   0.527    -.0018855    .0036824
          wi |   .0022548   .0003321     6.79   0.000     .0016039    .0029058
          cy |   .0001122   9.38e-06    11.96   0.000     .0000938    .0001306
          le |   .0022507   .0000774    29.09   0.000     .0020991    .0024024
          he |   -.003132   .0002753   -11.38   0.000    -.0036715   -.0025924
      ln_pop |    .406027     .02019    20.11   0.000     .3664511    .4456028
      ln_gdp |   .1272735   .0051288    24.82   0.000     .1172202    .1373267
       _cons |   -5.51132   .4082338   -13.50   0.000    -6.311528   -4.711112
------------------------------------------------------------------------------

Absorbed degrees of freedom:
-----------------------------------------------------+
 Absorbed FE | Categories  - Redundant  = Num. Coefs |
-------------+---------------------------------------|
          ma |         5           0           5     |
          ye |        30           1          29     |
         brd |        40           1          39    ?|
-----------------------------------------------------+
? = number of redundant parameters may be higher
```

Given these parameter estimates, we can estimate marginal cost as the fitted value (including fixed effects) of the previous regression using the "predict *varname*, xbd" command. Since marginal revenue is in logarithms, we will have to take the exponential of the fitted value in order to obtain marginal cost in euros.

```
. predict ln_mc, xbd

. gen mc = exp(ln_mc)
```

Finally, taking the average of the marginal costs we have generated across firms, we obtain the following histogram:

```
. bysort frm: egen mean_mc = mean(mc)

. hist mean_mc, bin(25)
(bin=25, start=259.58572, width=77.513759)
```