**EMPIRICAL INDUSTRIAL ORGANIZATION (ECO 310)**

**Fall 2022 – Victor Aguirregabiria**

**SOLUTION TO PROBLEM SET 2**

**November 20, 2022**

---

**INSTRUCTIONS.** Please, follow these instructions for the submission of your completed problem set.

1. Write your answers electronically in a word processor.
2. For the answers that involve coding in STATA, include in the document the code in STATA that you have used to obtain your empirical results.
3. Convert the document to PDF format.
4. Submit your problem set in PDF online via Quercus.
5. You should submit your completed problem set by **Sunday, November 20, 11:59pm**.
6. You can discuss about the problem set with you classmates, but your answers and code should be written individually.

**The total number of marks is 200.**

---

To answer the Questions in this Problem Set, you need to use the datafile `datafile_problemset_02_2022.dta` that you can download from the course website in Quercus. Use this dataset to implement the estimations described below. Please, provide the STATA code that you use to obtain the results.

This dataset contains information on the retail wine industry in a Canadian province. It is a panel dataset with three dimensions: wine product, retail store, and month. It includes $11,033$ wine products, 623 stores, and 12 months, and a total of $6,180,915$ product-store-month observations.

The following Table provides a brief description of all the variables in this dataset.

| Description of datafile | |
|---|---|
| `datafile_problemset_02_2022.dta` | |
| **Variable name** | **Description** |
| *product* | Wine Product ID Number |
| *store* | Store ID Number |
| *period* | Month count: from 1 to 12 |
| *qunit* | Quantity (750ml bottles) sold of product in store and month |
| *price_750ml* | Price of Product per 750ml bottle |
| *alc* | Alcohol percentage points of Wine Product |
| *sugar_gpl* | Sugar (gram per litter) of Wine Product |
| *redwine* | Dummy for Red Wine |
| *whitewine* | Dummy for White Wine |
| *num_country* | Country of origin of Wine Product |
| *ontario* | Dummy for Ontario origin of Wine Product |
| *winerack* | Dummy for Wine Product belongs Wine Rack Brand |
| *wineshop* | Dummy for Wine Product belongs Wine Shop Brand |
| *otherontario* | Dummy for Ontario Wine other than Wine Rack or Wine Shop |
| *local_msize* | Market size (in bottles of wine) at store-month |

For the rest of this problem set, we use the following subindexes: $j$ for product, $m$ for store (local market), and $t$ for month.

Consider the following Logit Demand Model:

$$\ln\left(\frac{s_{jmt}}{s_{0mt}}\right) = -\alpha\, p_{jt} + \mathbf{x}_j\, \boldsymbol{\beta} + \xi_{jmt} \tag{1}$$

Variable $s_{jmt}$ is the market share of product $j$ in store $m$ and month $t$, that is, $s_{jmt} = \frac{qunit}{localmsize}$. Variable $p_{jt}$ represents the price of product $j$ at month $t$, that is, $p_{jt} = price750ml$. Note that, in this province, the price of a product is the same across all the local markets (stores). Finally, vector $\mathbf{x}_j$ contains the following product characteristics:

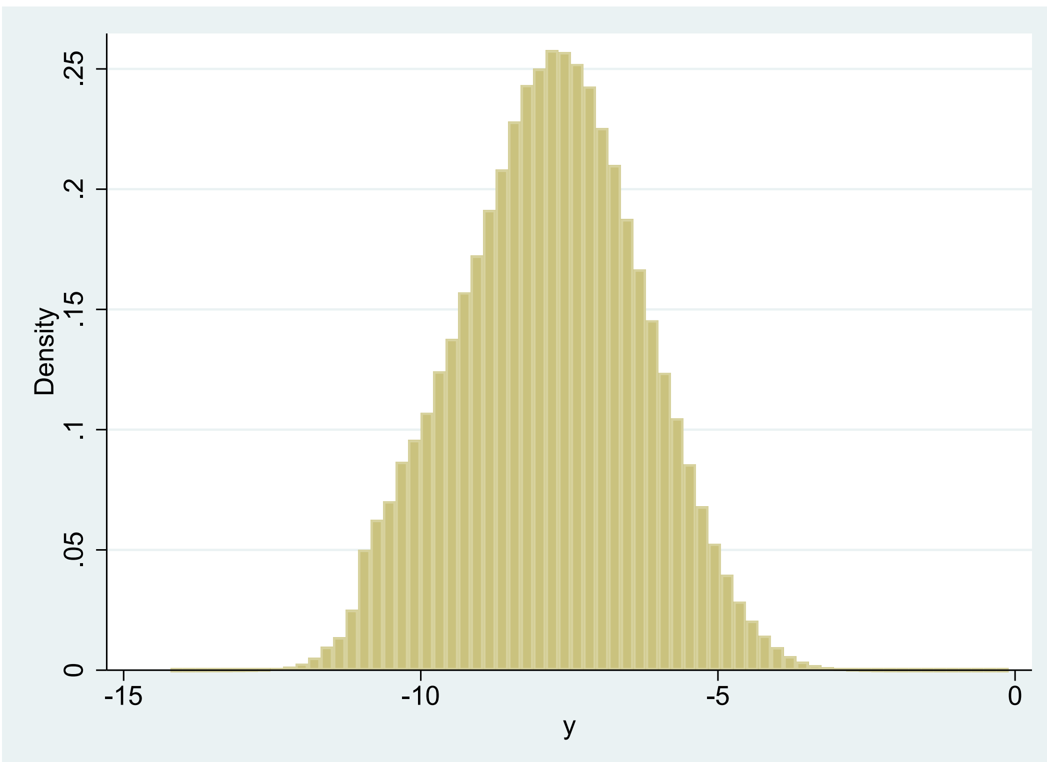
$\mathbf{x}_j = (alc,\ sugargpl,\ redwine,\ whitewine,\ winerack,\ wineshop,\ otherontario,\ Country\ dummies)$

**QUESTION 1. [20 points].** (A) Construct market shares $s_{jmt}$ and $s_{0mt}$, and variable $y_{jmt} = \ln(s_{jmt}/s_{0mt})$. (B) Present a figure with the histogram of $y$.

ANSWER: These are the lines of code in STATA.

```
gen share = qunit/local_msize
egen sumshare = sum(share), by(store period)
gen share0 = 1 - sumshare
gen y = ln(share)-ln(share0)
hist y
```

This is the figure with the histogram of $y$.




**QUESTION 2. [20 points].** Obtain the OLS estimates of parameters $\alpha$ and $\boldsymbol{\beta}$ in equation (1). In this regression, include store fixed effects and month fixed effects. When reporting your table of estimation results, please do not include estimated coefficients for store dummies and month dummies.

ANSWER: These are the lines of code in STATA.

```
reghdfe y price_750ml alc sugar_gpl redwine whitewine winerack wineshop otherontario
i.num_country, a(store period)
gen alpha_q2 = -_b[price_750ml]
```

This is the table of estimation results.

```
. reghdfe y price_750ml alc sugar_gpl redwine whitewine winerack wineshop otherontario i.num_country, a(store period)
(MWFE estimator converged in 4 iterations)

HDFE Linear regression                            Number of obs   =  5,234,233
Absorbing 2 HDFE groups                           F(  39,5233560) =   39136.46
                                                  Prob > F        =     0.0000
                                                  R-squared       =     0.4251
                                                  Adj R-squared   =     0.4250
                                                  Within R-sq.    =     0.2258
                                                  Root MSE        =     1.1166
```

| y | Coefficient | Std. err. | t | P>\|t\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| price_750ml | -.0400237 | .0000443 | -904.15 | 0.000 | -.0401104 | -.0399369 |
| alc | -.0718711 | .0002776 | -258.91 | 0.000 | -.0724152 | -.0713271 |
| sugar_gpl | .0018357 | .000021 | 87.41 | 0.000 | .0017946 | .0018769 |
| redwine | .2939236 | .0019097 | 153.91 | 0.000 | .2901807 | .2976665 |
| whitewine | .2419486 | .0019394 | 124.76 | 0.000 | .2381475 | .2457497 |
| winerack | .0059154 | .0034163 | 1.73 | 0.083 | -.0007804 | .0126112 |
| wineshop | .0889334 | .0035737 | 24.89 | 0.000 | .0819291 | .0959377 |
| otherontario | -.8095651 | .0031035 | -260.85 | 0.000 | -.8156478 | -.8034823 |
| num_country | | | | | | |
| Australia | -.0493464 | .0029291 | -16.85 | 0.000 | -.0550874 | -.0436054 |
| Austria | -1.663765 | .0154882 | -107.42 | 0.000 | -1.694122 | -1.633409 |
| Bulgaria | -.5142011 | .0243035 | -21.16 | 0.000 | -.5618351 | -.466567 |
| Canada | .0435827 | .0037798 | 11.53 | 0.000 | .0361745 | .0509909 |
| Chile | -.0138141 | .0032518 | -4.25 | 0.000 | -.0201875 | -.0074407 |
| France | -.2946378 | .0029452 | -100.04 | 0.000 | -.3004103 | -.2888653 |
| Georgia | -1.088462 | .0398413 | -27.32 | 0.000 | -1.16655 | -1.010375 |
| Germany | -.2864667 | .0045466 | -63.01 | 0.000 | -.2953779 | -.2775556 |
| Greece | -.8517643 | .0087705 | -97.12 | 0.000 | -.8689542 | -.8345743 |
| Hungary | -.4024938 | .0091775 | -43.86 | 0.000 | -.4204814 | -.3845061 |
| Israel | -.9708158 | .0179022 | -54.23 | 0.000 | -1.005903 | -.9357282 |
| Italy | -.0251114 | .0028027 | -8.96 | 0.000 | -.0306045 | -.0196182 |
| Jamaica | -.7160749 | .0287141 | -24.94 | 0.000 | -.7723536 | -.6597963 |
| Japan | -.8304819 | .0094734 | -87.66 | 0.000 | -.8490494 | -.8119145 |
| Lebanon | -1.922526 | .0761997 | -25.23 | 0.000 | -2.071875 | -1.773177 |
| Luxembourg | -.7329813 | .0714058 | -10.27 | 0.000 | -.8729343 | -.5930284 |
| Mexico | -.5580269 | .0245101 | -22.77 | 0.000 | -.6060659 | -.5099879 |
| Montenegro | -.8259244 | .0333839 | -24.74 | 0.000 | -.8913557 | -.7604932 |
| New Zealand | -.2916938 | .0044912 | -64.95 | 0.000 | -.3004963 | -.2828912 |
| Poland | -.5890749 | .0270352 | -21.79 | 0.000 | -.642063 | -.5360868 |
| Portugal | -.3176446 | .004158 | -76.39 | 0.000 | -.3257941 | -.309495 |
| Republic of Macedonia | -.8692797 | .0325435 | -26.71 | 0.000 | -.9330638 | -.8054956 |
| Romania | -.4582615 | .0243662 | -18.81 | 0.000 | -.5060183 | -.4105047 |
| Serbia | -.4705951 | .0163173 | -28.84 | 0.000 | -.5025765 | -.4386137 |
| South Africa | -.1961989 | .0038493 | -50.97 | 0.000 | -.2037435 | -.1886543 |
| South Korea | -1.204587 | .0321621 | -37.45 | 0.000 | -1.267623 | -1.14155 |
| Spain | -.4254888 | .0038105 | -111.66 | 0.000 | -.4329572 | -.4180205 |
| Switzerland | -1.501503 | .1183965 | -12.68 | 0.000 | -1.733556 | -1.26945 |
| USA | -.0059063 | .0029572 | -2.00 | 0.046 | -.0117023 | -.0001103 |
| Ukraine | -1.00718 | .0794732 | -12.67 | 0.000 | -1.162945 | -.8514152 |
| United Kingdom | -.2732882 | .0320088 | -8.54 | 0.000 | -.3360242 | -.2105521 |
| _cons | -6.193492 | .0045431 | -1363.27 | 0.000 | -6.202396 | -6.184588 |

**QUESTION 3. [20 points].** Based on your estimates in Question 2, provide an estimate of the willingness to pay (in dollars per 750ml bottle) of the average consumer for the following changes.

ANSWER: Let me start providing a general description of the calculation of a consumer maximum willingness to pay (WTP) for switching from a product with characteristics $\mathbf{x}_j$ to other product with characteristics $\mathbf{x}_k$. The WTP is equal to the change in price that leaves the consumer indifferent (same utility) between buying product $j$ and product $k$. That is, the equation that defines the WTP is:

$$\mathbf{x}_j \boldsymbol{\beta} - \alpha \, p_j \; = \; \mathbf{x}_k \boldsymbol{\beta} - \alpha \, \left( p_j + WTP \right)$$

Solving for WTP, we have:

$$WTP \; = \; \frac{\left( \mathbf{x}_k - \mathbf{x}_j \right) \boldsymbol{\beta}}{\alpha}$$

For instance, if the only difference between $\mathbf{x}_k$ and $\mathbf{x}_j$ is that $k$ is red wine and $j$ is white wine, we have:

$$WTP \; = \; \frac{\beta_{redwine} - \beta_{whitewine}}{\alpha}$$

a. A reduction in alcohol content of 1 percent point.

   ANSWER: This is equal to $\frac{-\beta_{alc}}{\alpha}$. Note that it is a REDUCTION in alcohol content, and this is why the negative sign multiplying $\beta_{alc}$. Note also that $\alpha = -\_b[price]$. The line of code in STATA is:

```
dis (-_b[alc])/(-_b[price])
```

   And the result is:

```
1.7957159
```

   The average consumer is willing to pay $1.79 more per bottle if the alcohol content is reduced in 1 percent point.

b. An increase in sugar content of 1 gram per liter.

   ANSWER: This is equal to $\frac{\beta_{sugar\_gpl}}{\alpha}$. The line of code in STATA is:

```
dis _b[sugar_gpl]/(-_b[price])
```

And the result is:

```
0.04586609
```

The average consumer is willing to pay only 4 cents more per bottle if the sugar content increases in 1 gram per liter.

c. Switching from a white wine to a read wine.

ANSWER: This is equal to $\frac{\beta_{redwine}-\beta_{whitewine}}{\alpha}$. The line of code in STATA is:

```
dis (_b[redwine]-_b[whitewine])/(-_b[price])
```

And the result is:

```
1.2986066
```

The average consumer is willing to pay $1.29 more for a bottle of red wine than for a bottle of white wine.

d. Switching from ″otherontario″ wine to ″wineshop″ wine.

ANSWER: This is equal to $\frac{\beta_{wineshop}-\beta_{otherontario}}{\alpha}$. The line of code in STATA is:

```
dis (_b[wineshop]-_b[otherontario])/(-_b[price])
```

And the result is:

```
22.449176
```

The average consumer is willing to pay $22.44 more for a bottle of Wineshop wine than for a bottle of otherontario wine. The average consumer perceives a very large difference in quality between Ontario wines of the WineShop brand and Ontario wines that do not belong to WineShop or WineRack brands (otherontario).

e. Switching from a French wine to a Canadian wine.

ANSWER: This is equal to $\frac{\beta_{Canada} - \beta_{France}}{\alpha}$. To obtain this WTP in STATA, it is important to see that "Canada" corresponds to $num\_country == 6$, and "France" corresponds to $num\_country == 9$. The line of code in STATA is:

```
dis (_b[6.num_country]-_b[9.num_country])/(-_b[price])
```

And the result is:

```
8.4505123
```

The average consumer is willing to pay $8.45 more for a bottle of Canadian wine than for a bottle of French wine. It could be interpreted in terms of average quality differences, but also in terms of preference bias towards domestic products.

**QUESTION 4. [10 points].** Based on the estimates in Question 2, obtain a variable with the estimated own-price demand elasticity $-\frac{ds_{jmt}}{dp_{jmt}} \frac{p_{jmt}}{s_{jmt}}$ for every observation $(j, m, t)$ in the data. Present the mean and median of this variable, and a figure with its histogram.

ANSWER: For the logit model, the derivative $\frac{ds_{jmt}}{dp_{jmt}}$ is equal to $-\alpha s_{jmt}(1 - s_{jmt})$. Therefore, the own-price demand elasticity $-\frac{ds_{jmt}}{dp_{jmt}} \frac{p_{jmt}}{s_{jmt}}$ is equal to $\alpha p_{jmt}(1 - s_{jmt})$. The lines of code in STATA are:

```
gen elast_q2 = alpha_q2 * (1-share) * price_750ml
sum elast_q2, detail // Summary statistics
hist elast_q2 // Histogram
```

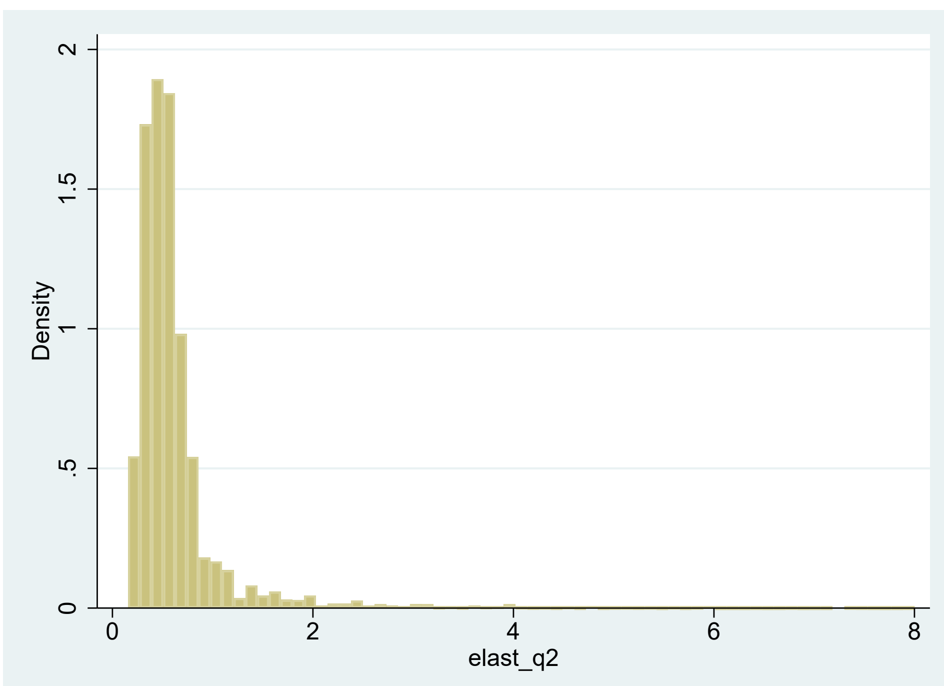
This is the table of Summary Statistics.

```
. sum elast_q2, detail     // Summary statistics

                          elast_q2
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .2390613       .1561722
 5%     .2630428       .1580908
10%     .2918696       .1590348       Obs           6,180,915
25%     .3780208       .1605566       Sum of wgt.   6,180,915

50%     .5178548                      Mean           .6209684
                        Largest       Std. dev.      .5163205
75%       .66789       7.984651
90%     .9584959       7.994168       Variance       .2665869
95%     1.402332       8.002663       Skewness       5.020582
99%     3.001732       8.002674       Kurtosis       39.68019
```

The mean elasticity is 0.62 and the median elasticity is $0,51$. This is the Histogram of estimates elasticities.



**QUESTION 5. [20 points].** Suppose that $\xi_{jmt}$ follows an AR(1) process: $\xi_{jmt} = \rho\, \xi_{jm,t-1} + a_{jmt}$. Obtain the OLS Cochrane-Orcutt estimates of parameters $\rho$, $\alpha$, $\rho\alpha$, and $(1-\rho)\boldsymbol{\beta}$. In this regression, include store fixed effects and month fixed effects. Do not report coefficients for store dummies and month dummies.

ANSWER: The Cochrane-Orcutt transformation of the model is a "Quasi-first-difference": we multiply the equation at period $t-1$ times $\rho$ and substract this equation to the equation at period $t$. We obtain the following regression equation:

$$y_{jmt} \;=\; \gamma_1\, y_{jm,t-1} \;+\; \gamma_2\, p_{jmt} \;+\; \gamma_3\, p_{jm,t-1} \;+\; \mathbf{x}_j\, \boldsymbol{\gamma_4} \;+\; a_{jmt}$$

where $\gamma_1 = \rho$, $\gamma_2 = -\alpha$, $\gamma_3 = \rho\alpha$, and $\boldsymbol{\gamma_4} = (1-\rho)\boldsymbol{\beta}$. We apply OLS with Fixed Effects to this regression equation. These are the lines of Code in STATA.

```
// Construction of lagged y
sort product store period
gen y_1 = y[_n-1] if (product==product[_n-1])  (store==store[_n-1])

// Construction of lagged price
gen price_1 = price[_n-1] if (product==product[_n-1])  (store==store[_n-1])

// OLS - Cochrane-Orcutt regression
reghdfe y y_1 price_750ml price_1 alc sugar_gpl redwine whitewine winerack wineshop
otherontario i.num_country, a(store period)

gen alpha_q5 = -_b[price_750ml]
```

This is the Table with the Cochrane-Orcutt FE estimates.

```
. reghdfe y y_1 price_750ml price_1 alc sugar_gpl redwine whitewine winerack wineshop otherontario i.num_country, a(store period)
(MWFE estimator converged in 4 iterations)
```

HDFE Linear regression
Absorbing 2 HDFE groups

| | | |
|---|---|---|
| Number of obs | = | 4,524,065 |
| F( 41,4523391) | = | 194654.85 |
| Prob > F | = | 0.0000 |
| R-squared | = | 0.7314 |
| Adj R-squared | = | 0.7314 |
| Within R-sq. | = | 0.6383 |
| Root MSE | = | 0.7553 |

| y | Coefficient | Std. err. | t | P>\|t\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| y_1 | .7328674 | .0003234 | 2266.08 | 0.000 | .7322335 | .7335012 |
| price_750ml | -.2354997 | .0005655 | -416.45 | 0.000 | -.236608 | -.2343913 |
| price_1 | .2244183 | .0005674 | 395.51 | 0.000 | .2233062 | .2255304 |
| alc | -.0191792 | .0002017 | -95.08 | 0.000 | -.0195746 | -.0187838 |
| sugar_gpl | .0004285 | .0000154 | 27.85 | 0.000 | .0003983 | .0004586 |
| redwine | .0626643 | .001406 | 44.57 | 0.000 | .0599085 | .0654201 |
| whitewine | .0633322 | .001427 | 44.38 | 0.000 | .0605353 | .0661291 |
| winerack | .0128959 | .0024634 | 5.23 | 0.000 | .0080677 | .0177242 |
| wineshop | .0293741 | .0025705 | 11.43 | 0.000 | .0243361 | .0344122 |
| otherontario | -.214636 | .0022605 | -94.95 | 0.000 | -.2190666 | -.2102055 |
| num_country | | | | | | |
| Australia | -.0258695 | .0021431 | -12.07 | 0.000 | -.0300699 | -.0216692 |
| Austria | -.468588 | .0117506 | -39.88 | 0.000 | -.4916188 | -.4455572 |
| Bulgaria | -.1408248 | .0173303 | -8.13 | 0.000 | -.1747916 | -.1068581 |
| Canada | -.0047509 | .0027476 | -1.73 | 0.084 | -.0101361 | .0006344 |
| Chile | -.009832 | .00237 | -4.15 | 0.000 | -.0144772 | -.0051869 |
| France | -.0891188 | .0021644 | -41.17 | 0.000 | -.093361 | -.0848766 |
| Georgia | -.3444855 | .0362721 | -9.50 | 0.000 | -.4155775 | -.2733935 |
| Germany | -.0884129 | .0032878 | -26.89 | 0.000 | -.0948569 | -.0819689 |
| Greece | -.2353675 | .0065117 | -36.15 | 0.000 | -.2481303 | -.2226047 |
| Hungary | -.1111851 | .0066636 | -16.69 | 0.000 | -.1242455 | -.0981247 |
| Israel | -.1852285 | .0139085 | -13.32 | 0.000 | -.2124886 | -.1579683 |
| Italy | -.0098927 | .0020529 | -4.82 | 0.000 | -.0139163 | -.0058692 |
| Jamaica | -.2098928 | .0205262 | -10.23 | 0.000 | -.2501234 | -.1696621 |
| Japan | -.2330154 | .0068837 | -33.85 | 0.000 | -.2465072 | -.2195236 |
| Lebanon | -.4351287 | .0727157 | -5.98 | 0.000 | -.5776489 | -.2926086 |
| Luxembourg | -.2092075 | .0619266 | -3.38 | 0.001 | -.3305814 | -.0878337 |
| Mexico | -.2236182 | .0189948 | -11.77 | 0.000 | -.2608473 | -.186389 |
| Montenegro | -.2488628 | .0238747 | -10.42 | 0.000 | -.2956563 | -.2020693 |
| New Zealand | -.0738408 | .0033309 | -22.17 | 0.000 | -.0803693 | -.0673123 |
| Poland | -.1746559 | .0193909 | -9.01 | 0.000 | -.2126613 | -.1366506 |
| Portugal | -.0875834 | .0030301 | -28.90 | 0.000 | -.0935223 | -.0816445 |
| Republic of Macedonia | -.227967 | .0232284 | -9.81 | 0.000 | -.2734939 | -.18244 |
| Romania | -.1661927 | .0182541 | -9.10 | 0.000 | -.20197 | -.1304153 |
| Serbia | -.1395491 | .0116603 | -11.97 | 0.000 | -.1624028 | -.1166953 |
| South Africa | -.0488728 | .002813 | -17.37 | 0.000 | -.0543862 | -.0433594 |
| South Korea | -.3392482 | .0236212 | -14.36 | 0.000 | -.3855449 | -.2929515 |
| Spain | -.1269013 | .0028143 | -45.09 | 0.000 | -.1324173 | -.1213853 |
| Switzerland | -.5430372 | .1315132 | -4.13 | 0.000 | -.8007983 | -.2852761 |
| USA | -.0016444 | .002173 | -0.76 | 0.449 | -.0059033 | .0026146 |
| Ukraine | -.4844619 | .0720937 | -6.72 | 0.000 | -.6257631 | -.3431607 |
| United Kingdom | -.0981158 | .0229296 | -4.28 | 0.000 | -.143057 | -.0531745 |
| _cons | -1.626884 | .0038495 | -422.62 | 0.000 | -1.634429 | -1.619339 |

COMMENTS ON Cochrane-Orcutt FE estimates.

- $\rho = \_b[y_1] = 0.7328$. There is substantial serial correlation in the demand unobservables $\xi_{jmt}$.
- $\alpha = -\_b[price\_750ml] = 0.2354$. This estimate of the price coefficient is six times larger the OLS-FE in Question 2 (0.040). This is consistent with strong simultaneity bias in the OLS-FE in Question 2 due to correlation between $p_{jmt}$ and $\xi_{jmt}$, and more specifically, due to correlation between $p_{jmt}$ and $\rho\ \xi_{jm,t-1}$. By including $y_{jm,t-1}$ and $p_{jm,t-1}$ as regressors, the Cochrane-Orcutt FE estimator is controlling for $\rho\ \xi_{jm,t-1}$, and therefore, it does not suffer of bias because correlation between $p_{jmt}$ and $\rho\ \xi_{jm,t-1}$. The Cochrane-Orcutt FE estimator can be still biased because correlation between $p_{jmt}$ and $a_{jmt}$, but this source of bias is also present in the OLS-FE estimator in Question 2. There is a solid argument for Cochrane-Orcutt FE estimator to be substantially less biased than OLS-FE.
- Cochrane-Orcutt model implies the over-identifying restriction: $\gamma_1\gamma_2 + \gamma_3 = 0$. The estimate of $\gamma_1\gamma_2 + \gamma_3$ (`display _b[y_1]*_b[price_750ml] +_b[price_1]`) is equal to 0.0518, which is "close to zero", given the magnitude of the parameter estimates. However, with this large sample, parameter estimates are very precise, and the test of this over-identifying restriction (`testnl _b[y_1]*_b[price_750ml] +_b[price_1] = 0`) clearly resjects it, with a p-value smaller than 0.01%.
- To compare the estimates in Questions 2 and 5 of parameters for product characteristics other than price (the variable in vector $\mathbf{x}_j$). we need to take into account that $\gamma = (1-\rho)\beta$. For instance, in Question 5, $\gamma_{alc} = -0.0191$, that implies $\beta_{alc} = \gamma_{alc}/(1-\rho) = -0.0191/(1-0.7328) = -0.07179$. This is almost identical to the estimate of $\beta_{alc}$ in Question 2 which is $-0.07187$.

**QUESTION 6. [10 points].** Based on the estimates in Question 5, obtain a variable with the estimated own-price demand elasticity $-\frac{ds_{jmt}}{dp_{jmt}} \frac{p_{jmt}}{s_{jmt}}$ for every observation $(j, m, t)$ in the data. Present the mean and median of this variable, and a figure with its histogram.

ANSWER: For the logit model, the derivative $\frac{ds_{jmt}}{dp_{jmt}}$ is equal to $-\alpha s_{jmt}(1 - s_{jmt})$. Therefore, the own-price demand elasticity $-\frac{ds_{jmt}}{dp_{jmt}} \frac{p_{jmt}}{s_{jmt}}$ is equal to $\alpha p_{jmt}(1 - s_{jmt})$. The lines of code in STATA are:

```
gen elast_q5 = alpha_q5 * (1-share) * price_750ml
sum elast_q5, detail // Summary statistics
hist elast_q5 // Histogram
```

This is the table of Summary Statistics.

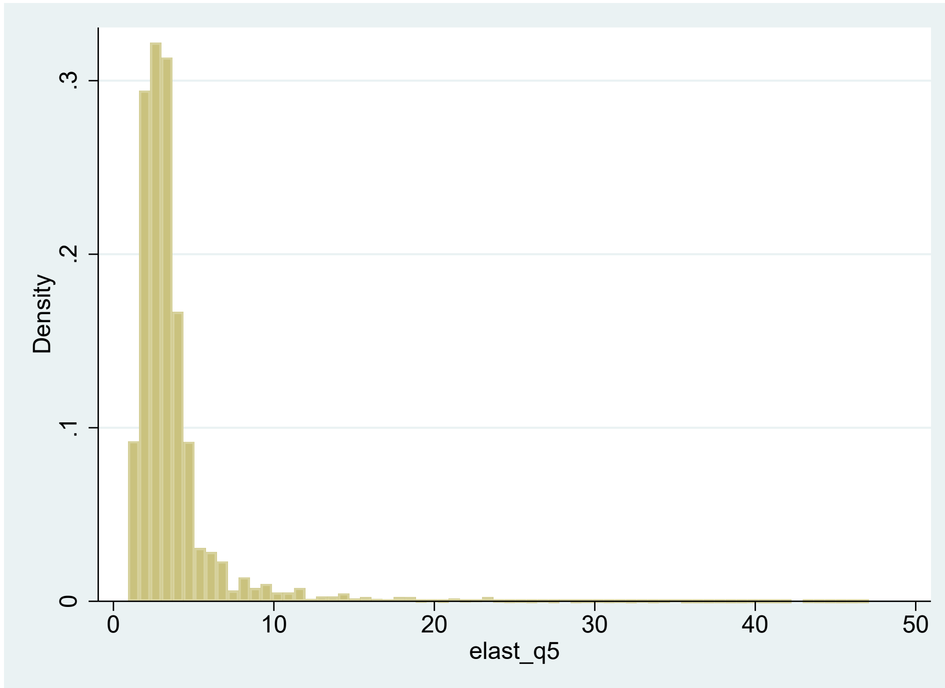
```
. sum elast_q5, detail    // Summary statistics

                          elast_q5
-------------------------------------------------------------
      Percentiles      Smallest
 1%     1.406639       .9189186
 5%     1.547747        .930208
10%     1.717363       .9357624       Obs           6,180,915
25%     2.224278       .9447165       Sum of wgt.   6,180,915

50%     3.047062                      Mean           3.653784
                        Largest       Std. dev.      3.038035
75%     3.929871       46.98177
90%       5.6398       47.03777       Variance       9.229658
95%     8.251336       47.08775       Skewness       5.020582
99%     17.66222       47.08782       Kurtosis       39.68019
```

Now, the mean elasticity is 3.65 and the median elasticity is 3.04. These elasticities are around six times the ones in Question 4.

This is the histogram of elasticities.




**QUESTION 7. [10 points].** Taking into account potential endogeneity, interpret the different estimates of $\alpha$ and of the corresponding demand elasticities in Questions 2 and 5. Based on this interpretation, select between one of these two demand estimations. Justify your choice.

ANSWER:

The CO-FE estimates of $\alpha$ and of price elasticities in Question 5 are six times larger the OLS-FE estimates in Question 2. This is consistent with strong simultaneity bias in the OLS-FE in Question 2 due to correlation between $p_{jmt}$ and $\xi_{jmt}$, and more specifically, due to correlation between $p_{jmt}$ and $\rho\ \xi_{jm,t-1}$. By including $y_{jm,t-1}$ and $p_{jm,t-1}$ as regressors, the Cochrane-Orcutt FE estimator is controlling for $\rho\ \xi_{jm,t-1}$, and therefore, it does not suffer of bias because correlation between $p_{jmt}$ and $\rho\ \xi_{jm,t-1}$. The Cochrane-Orcutt FE estimator can be still biased because correlation between $p_{jmt}$ and $a_{jmt}$, but this source of bias is also present in the OLS-FE estimator in Question 2. There is a solid argument for Cochrane-Orcutt FE estimator to be substantially less biased than OLS-FE. Between these two estimators, I would choose CO-FE.

---

For the remaining questions, we consider the following assumptions.

A. Each wine product is produced by a single manufacturer, and that each manufacturer produces only one wine product. Therefore, $j$ indexes both products and firms.

B. Firms (i.e., wine manufacturers) compete in prices a la Nash-Bertrand.

C. Each store $m$ is a separate geographic market where firms (i.e., wine manufacturers) compete with each other.

D. Every firm can choose a different for its wine product at each store.

E. The marginal cost of product $j$ in market $m$ at month $t$, $MC_{jmt}$, is constant: that is, $MC_{jmt}$ does not depend on the amount of output $q_{jmt}$.

**QUESTION 8. [10 points].**

a. Obtain the expression for the first order condition of profit maximization for firm $j$ in market $m$ at month $t$.

ANSWER: Under conditions A to E, the profit of firm $j$ in store $m$ at period $t$ is:

$$\pi_{jmt} = H_{mt} \, \left(p_{jmt} - MC_{jmt}\right) \, s_{jmt}$$

where $H_{mt}$ represent market size (number of consumers) in $m$ at period $t$. Under Nash-Bertrand competition, the first order condition for profit maximization is:

$$\frac{d\pi_{jmt}}{dp_{jmt}} \;=\; H_{mt} \, s_{jmt} \;+\; H_{mt} \, \left(p_{jmt} - MC_{jmt}\right) \, \frac{ds_{jmt}}{dp_{jmt}} \;=\; 0$$

For the logit demand model, $\frac{ds_{jmt}}{dp_{jmt}} = -\alpha \, s_{jmt}(1 - s_{jmt})$. Pluging this expression into the equation for the marginal condition for profit maximization, we get:

$$H_{mt} \, s_{jmt} \;-\; H_{mt} \, \left(p_{jmt} - MC_{jmt}\right) \, \alpha \, s_{jmt} \, (1 - s_{jmt}) \;=\; 0$$

b Based on this condition, obtain a "pricing equation" for price $p_{jmt}$ in terms of $MC_{jmt}$ and a price-cost margin term that depends only on $\alpha$ and $s_{jmt}$.

ANSWER: In the F.O.C. for profit maximization, we have divide both sides of the equation by $H_{mt} \, s_{jmt}$ to obtain:

$$1 \;-\; \left(p_{jmt} - MC_{jmt}\right) \, \alpha \, (1 - s_{jmt}) \;=\; 0$$

Solving for $p_{jmt}$, we get the pricing equation:

$$p_{jmt} \;=\; MC_{jmt} \;+\; \frac{1}{\alpha \, (1 - s_{jmt})}$$

The term $\dfrac{1}{\alpha \, (1 - s_{jmt})}$ represents the Price-Cost Margin in dollars.

c. Define the price-cost margin $PCM_{jmt}$ as $\frac{p_{jmt} - MC_{jmt}}{p_{jmt}}$. Obtain an expression for the price-cost margin in terms only of $\alpha$, $s_{jmt}$, and $p_{jmt}$.

ANSWER: $PCM_{jmt} \equiv \dfrac{p_{jmt} - MC_{jmt}}{p_{jmt}}$ is the Price-Cost Margin as a percentage of the price. It is also denoted the *Lerner Index*. Using the pricing equation in Question 8(b), we have that:

$$PCM_{jmt} \equiv \frac{p_{jmt} - MC_{jmt}}{p_{jmt}} = \frac{1}{\alpha \; p_{jmt} \; (1 - s_{jmt})}$$

**QUESTION 9. [20 points].** Use the expressions in Question 8 to obtain estimates of $MC_{jmt}$ and of $PCM_{jmt}$ for every observation $(j, m, t)$. Use your favorite estimate of $\alpha$. Obtain the mean and median of $PCM$. Present a histogram of the estimated $PCM$s.

ANSWER: These are the lines of code in STATA:

```
gen MCest = price_750ml - 1/(alpha_q5 * (1-share))
gen PCMest = (price_750ml - MCest)/price_750ml
sum PCMest, detail
hist PCMest
```

This is the table of summary statistics for PCMest:

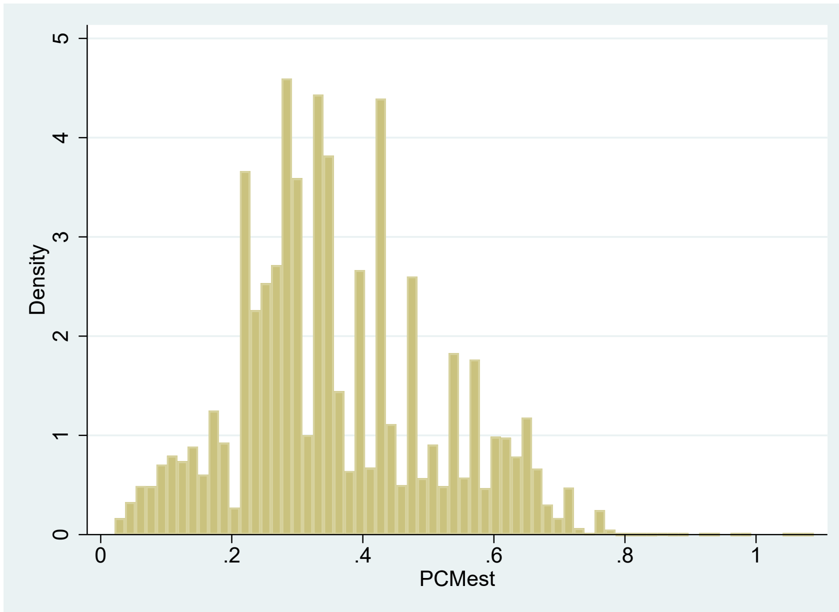
```
. sum PCMest, detail

                           PCMest
-------------------------------------------------------------

      Percentiles      Smallest
 1%      .056618       .0212369
 5%     .1211925       .0212369
10%     .1773113       .0212595       Obs           6,180,915
25%     .2544613       .0212848       Sum of wgt.   6,180,915

50%      .328185                      Mean           .3588846
                        Largest       Std. dev.      .1512186
75%     .4495841       1.058519
90%     .5822879       1.068647       Variance       .0228671
95%     .6461005       1.075028       Skewness       .3753284
99%     .7109144       1.088236       Kurtosis       2.682985
```

The mean PCM is 35.8%, and the median is 32.8%. This is the Histogram of PCMs:




---

For Question 10 below, take into account the equation for market share in a demand logit model:

$$s_{jmt} = \frac{a_{jmt} \exp\{-\alpha \, p_{jmt} + w_{jmt}\}}{1 + \sum_{i=1}^{J} a_{imt} \exp\{-\alpha \, p_{imt} + w_{imt}\}} \tag{2}$$

where $a_{jmt} \in \{0,1\}$ is a binary variable that indicates whether product $j$ is available in store $m$ at month $t$. i.e., $a_{jmt}$ is the indicator of $q_{jmt} > 0$. Variable $w_{jmt}$ is equal to $\mathbf{x}_j \, \boldsymbol{\beta} + \xi_{jmt}$, and it can be interpreted as an index of product quality. Also, note that in the logit demand model, taking into account equation (1), we can obtain $w_{jmt}$ using the following equation:

$$w_{jmt} = \ln\left(\frac{s_{jmt}}{s_{0mt}}\right) + \alpha \, p_{jt} \tag{3}$$

**QUESTION 10. [60 points].** Given your estimates of demand parameters and MCs, we are interested in evaluating the effects on quantities, prices, firms' profits, and consumer welfare of a hypothetical (counterfactual) increase in the tax of alcohol. Suppose that the tax increase is equal to 1 cent of a dollar per percentage of alcohol in the product, i.e., $taxchange_j = \$0.01 \, * \, alc_j$.

a. **[5 points]** Calculate new variable `taxchange` with $taxchange_{jmt}$ for every observation $(j, m, t)$. Add `taxchange` to the estimated marginal cost to obtain the new (counterfactual) marginal cost for every observation $(j, m, t)$.

ANSWER: These are the lines of code in StATA.

```
// Calculate taxchange
gen taxchange = 0.01 * alc

// Calculate New Marginal Cost
gen newMC = MCest + taxchange
```

b. **[5 points]** Using equation (3), calculate a new variable `wquality` with $w_{jmt}$ for every observation $(j, m, t)$.

ANSWER: This is the line of code in StATA.

```
gen wquality = y + alpha_q5 * price_750ml
```

c. **[20 points]** Obtain an approximation to the new (counterfactual) equilibrium prices and market shares using the following iterative procedure.

Step 1. Use the pricing equation in Question 8(b) to obtain new prices, say $p_{jmt}^{ITER\,1}$, using the new MCs, your estimate of $\alpha$, and $s_{jmt}$ in the data.

Step 2. Use equation (2) for market shares, your estimates for product qualities $w_{jmt}$, and new prices $p_{jmt}^{ITER\,1}$ to calculate new market shares, say $s_{jmt}^{ITER\,1}$.

Step 3. Similar as Step 1 but using market shares $s_{jmt}^{ITER\,1}$ instead of the shares in the data. Let $p_{jmt}^{ITER\,2}$ be the prices you obtain after applying this iteration.

Step 4. Similar to Step 2 but using prices $p_{jmt}^{ITER\,2}$ instead of $p_{jmt}^{ITER\,1}$. Let $s_{jmt}^{ITER\,2}$ be the shares you obtain after applying this iteration.

ANSWER: These are the lines of code in STATA.

```
// Step 1
gen p_iter_1 = newMC + 1/(alpha_q5 * (1-share))

// Step 2
gen buff = exp(-alpha_q5 * p_iter_1 + wquality)
egen agg_buff = sum(buff), by(store period)
gen s_iter_1 = buff/(1 + agg_buff)
```

```
drop buff agg_buff

// Step 3
gen p_iter_2 = newMC + 1/(alpha_q5 * (1-s_iter_1))

// Step 4
gen buff = exp(-alpha_q5 * p_iter_2 + wquality)
egen agg_buff = sum(buff), by(store period)
gen s_iter_2 = buff/(1 + agg_buff)
drop buff agg_buff
```

d. **[30 points]** Suppose that $p_{jmt}^{ITER\,2}$ and $s_{jmt}^{ITER\,2}$ are prices and market shares after the implementation of this counterfactual tax change. Calculate the following statistics at the aggregate annual and province level:

- **Median price change.**

  ANSWER: These are the lines of code in STATA.

  ```
  // d.  Median Price Change
  gen price_change = p_iter_2 - price_750ml
  sum price_change, detail
  ```

  This is the table of summary statistics:

```
. sum price_change, detail

                        price_change
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0600023       .0480013
 5%     .0949984       .0498319
10%     .1119576       .0499992       Obs           6,180,915
25%     .1199932       .0499992       Sum of wgt.   6,180,915

50%      .129981                      Mean           .1275753
                        Largest       Std. dev.      .0185573
75%     .1349945       .4699974
90%     .1399879       .4699974       Variance       .0003444
95%     .1449966       .4699974       Skewness       .5621972
99%     .1997428       .4699974       Kurtosis       23.46936
```

The median price change is 0.13

COMMENTS: The median of taxchange is 0.13 and the median price change is 0.13. Therefore, at the median there is a complete pass-through of the tax to prices "as if" there were a perfectly competitive market. However, this is not the case over the whole distribution of prices.

– **Change in annual sales of wine in this province.**

ANSWER: These are the lines of code in STATA.

```
// d.  Change in annual sales of wine in this province
gen change_q = (s_iter_2 - share) * local_msize
egen agg_change_q = sum(change_q)
sum agg_change_q
```

This is the table of summary statistics:

```
. sum agg_change_q
```

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| agg_change_q | 6,180,915 | -2344767 | 0 | -2344767 | -2344767 |

The change in annual sales of wine is of $2,344,767$ bottle of wine. The annual sales before the tax was 161 million bottles, so this change represents a reduction of 1.4%.

– **Government annual revenue from this tax increase.**

ANSWER: These are the lines of code in STATA.

```
// d.  Government annual revenue from this tax increase.
gen gov_revenue = taxchange * s_iter_2 * local_msize
egen agg_gov_revenue = sum(gov_revenue)
sum agg_gov_revenue
```

This the table of summary statistics:

```
. sum agg_gov_revenue

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
agg_gov_re~e |  6,180,915    1.96e+07           0   1.96e+07   1.96e+07
```

The Government annual revenue from this tax increase is $19.6 million.

- **Change in total annual firms' profits in this province.**

ANSWER: These are the lines of code in STATA.

```
// d.  Change in total annual firms' profits in this province.
gen profit_data = (price_750ml - MCest)* share * local_msize
gen profit_tax = (p_iter_2 - newMC)* s_iter_2 * local_msize
gen change_profit = profit_tax - profit_data
sum agg_change_profit
```

This the table of summary statistics:

```
. sum agg_change_profit

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
agg_change~t |  6,180,915    -9995187           0   -9995187   -9995187
```

The change in total annual firms' profits is $10 million. This is almost half of the increase in government revenue. So, half of the government revenue is at the expense of firms profits.

– **Change in annual Consumer Surplus in this province.**

ANSWER: These are the lines of code in STATA.

```
// d.  Change in annual Consumer Surplus in this province
gen change_csurplus = - price_change * s_iter_2 * local_msize + 0.5 *
price_change * change_q
egen agg_change_csurplus = sum(change_csurplus)
sum agg_change_csurplus
```

This the table of summary statistics:

```
. sum agg_change_csurplus
```

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| agg_change~s | 6,180,915 | -1.98e+07 | 0 | -1.98e+07 | -1.98e+07 |

The change in in annual Consumer Surplus is $19.8 million. The reduction of consumer surplus is dramatic and it is even larger than the increase in government revenue.

COMMENTS:

– **Annual Deadweight Loss of this tax: Change in consumer surplus + Change in firms' profits + Change in government revenue.**

ANSWER: These are the lines of code in STATA.

```
// d.  Annual Deadweight Loss of this tax
// Change in consumer surplus + Change in firms' profits + Change in
government revenue
gen DWL = change_csurplus + change_profit + gov_revenue
egen agg_DWL = sum(DWL)
sum agg_DWL
```

This the table of summary statistics:

```
. sum agg_DWL

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
     agg_DWL |  6,180,915   -1.01e+07           0  -1.01e+07  -1.01e+07
```

The DWL is 0 million, which is almost half of the government revenue. That is, for every dollar of government revenue there is 50 cents of net loss in social welfare.

**Hint to obtain the change in Consumer Surplus.** You can approximate the change in consumer surplus applying the following formula for each observation $(j, m, t)$. Let $(p^{data}, q^{data})$ be price and quantity in the data, and let $(p^{tax}, q^{tax})$ be price and quantity after the tax. Define the change in price as $\Delta p \equiv p^{tax} - p^{data}$, and the change in quantity as $\Delta q \equiv q^{tax} - q^{data}$. Then, the change in consumer surplus can be obtained using the following triangular approximation:

$$\Delta CS \equiv CS^{tax} - CS^{data} = -\Delta p \; q^{tax} + \frac{1}{2} \Delta p \; \Delta q \tag{4}$$