

Empirical Industrial Organization (ECO 310)  
Fall 2019. Victor Aguirregabiria

Solution to Problem Set #2

Due on Thursday, December 5th, 2019 [before 11:59pm]

**INSTRUCTIONS.** Please, follow the following 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 before 11:59pm of Monday, December 5th, 2019.
6. Problem sets should be written individually.

The total number of marks is 150.

**QUESTION 1.** [70 points]. Consider an industry with a differentiated product. There are two firms in this industry, firms  $A$  and  $B$ . Each firm produces and sells two brands of the differentiated product: brands  $A1$  and  $A2$  are produced by firm  $A$ , and brands  $B1$  and  $B2$  by firm  $B$ . The demand system is a logit demand model, where consumers choose between five different alternatives:  $j = 0$ , represents the consumer decision of no purchasing any product; and  $j = A1$ ,  $j = A2$ ,  $j = B1$ , and  $j = B2$  represent the consumer purchase of product  $A1$ ,  $A2$ ,  $B1$ , and  $B2$ , respectively. The utility of no purchase ( $j = 0$ ) is zero. The utility of purchasing product  $j \in \{A1, A2, B1, B2\}$  is  $\beta x_j - \alpha p_j + \varepsilon_j$ , where the variables and parameters have the interpretation that we have seen in class. Variable  $x_j$  is a measure of the quality of product  $j$ , e.g., the number of stars of the product according to consumer ratings. Therefore, we have that  $\beta > 0$ . The random variables  $\varepsilon_1$  and  $\varepsilon_2$  are independently and identically distributed over consumers with a type I extreme value distribution, i.e., Logit model of demand. Let  $H$  be

the number of consumers in the market. Let  $s_0, s_{A1}, s_{A2}, s_{B1}$ , and  $s_{B2}$  be the market shares of the five choice alternatives, such that  $s_j$  represents the proportion of consumers choosing alternative  $j$  and  $s_0 + s_{A1} + s_{A2} + s_{B1} + s_{B2} = 1$ .

**Q1.1. (5 points)** Based on this model, write the equation for the market share  $s_{A1}$  as a function of the prices and the qualities  $x$ 's of all the products.

ANSWER. The logit model with two products and utilities  $\beta x_j - \alpha p_j + \varepsilon_j$  implies that the market share of product 1 is:

$$s_{A1} = \frac{\exp\{\beta x_{A1} - \alpha p_{A1}\}}{1 + \sum_{j \in \{A1, A2, B1, B2\}} \exp\{\beta x_j - \alpha p_j\}}$$

**Q1.2. (5 points)** Obtain the expression for the derivatives: (a)  $\frac{\partial s_j}{\partial p_j}$ ; and (b)  $\frac{\partial s_j}{\partial p_k}$  for  $j \neq k$ . Write the expression for these derivatives in terms only of the market shares  $s_j$  and  $s_k$  and the parameters of the model.

ANSWER. We can write  $s_j = \frac{\exp\{\delta_j\}}{1 + \Sigma}$  with  $\delta_j = \beta x_j - \alpha p_j$  and  $\Sigma = \exp\{\delta_{A1}\} + \exp\{\delta_{A2}\} + \exp\{\delta_{B1}\} + \exp\{\delta_{B2}\}$ . Note that:

$$\begin{aligned} \frac{\partial s_j}{\partial \delta_j} &= \frac{\exp\{\delta_j\} \Sigma}{\Sigma^2} - \frac{\exp\{\delta_j\} \exp\{\delta_j\}}{\Sigma^2} \\ &= s_j - (s_j)^2 = s_j(1 - s_j) \end{aligned}$$

And,

$$\begin{aligned} \frac{\partial s_j}{\partial \delta_k} &= -\frac{\exp\{\delta_j\} \exp\{\delta_k\}}{\Sigma^2} \\ &= -s_j s_k \end{aligned}$$

(a) Using the chain rule of derivation,  $\frac{\partial s_j}{\partial p_j} = \frac{\partial s_j}{\partial \delta_j} \frac{\partial \delta_j}{\partial p_j}$ . Note that  $\frac{\partial \delta_j}{\partial p_j} = -\alpha$ . And as shown above  $\frac{\partial s_j}{\partial \delta_j} = s_j(1 - s_j)$ . Therefore:

$$\frac{\partial s_j}{\partial p_j} = -\alpha s_j(1 - s_j)$$

(b) Using the chain rule of derivation,  $\frac{\partial s_j}{\partial p_k} = \frac{\partial s_j}{\partial \delta_k} \frac{\partial \delta_k}{\partial p_k}$ . Note that  $\frac{\partial \delta_k}{\partial p_k} = -\alpha$ . And as shown above  $\frac{\partial s_j}{\partial \delta_k} = -s_j s_k$ . Therefore:

$$\frac{\partial s_j}{\partial p_k} = \alpha s_j s_k$$

---

The profit function of firm  $A$  is  $\pi_A = p_{A1} q_{A1} + p_{A2} q_{A2} - c_{A1} q_{A1} - c_{A2} q_{A2} - FC(x_{A1}) - FC(x_{A2})$ , where:  $q_j$  is the quantity sold by firm  $j$  (i.e.,  $q_j = H s_j$ );  $c_j$  is the marginal cost of producing good  $j$ , that is assumed constant, i.e., linear cost function; and  $FC(x_j)$  is the fixed cost of producing a good with quality  $x_j$ .

**Q1.3. (20 points)** Suppose that firms take the qualities  $x$  of their products as given and compete in prices ala Bertrand.

(a) Obtain the two equations that describe the marginal condition of profit maximization of firm  $A$  in this Bertrand game. Write this equation taking into account the specific form of  $\frac{\partial s_j}{\partial p_j}$  and  $\frac{\partial s_j}{\partial p_k}$  in the Logit model.

ANSWER.

(a) The f.o.c. of profit maximization with respect to  $p_{A1}$  is:  $q_{A1} + p_{A1} \frac{\partial q_{A1}}{\partial p_{A1}} - c_{A1} \frac{\partial q_{A1}}{\partial p_{A1}} + p_{A2} \frac{\partial q_{A2}}{\partial p_{A1}} - c_{A2} \frac{\partial q_{A2}}{\partial p_{A1}} = 0$ . Since  $q_j = H s_j$ , we have that  $\frac{\partial q_{A1}}{\partial p_{A1}} = H \frac{\partial s_{A1}}{\partial p_{A1}}$ , and  $\frac{\partial q_{A2}}{\partial p_{A1}} = H \frac{\partial s_{A2}}{\partial p_{A1}}$ . Taking into account the expressions for  $\frac{\partial s_j}{\partial p_j}$  and  $\frac{\partial s_j}{\partial p_k}$  in Q1.2, we have that  $\frac{\partial q_{A1}}{\partial p_{A1}} = -H \alpha s_{A1}(1 - s_{A1}) = -\alpha q_{A1}(1 - s_{A1})$ , and  $\frac{\partial q_{A1}}{\partial p_{A2}} = H \alpha s_{A1} s_{A2} = \alpha q_{A1} s_{A2}$ . Therefore, this F.O.C. is:

$$q_{A1} - \alpha q_{A1} (p_{A1} - c_{A1}) (1 - s_{A1}) + \alpha q_{A1} (p_{A2} - c_{A2}) s_{A2} = 0$$

Or equivalently,

$$\frac{1}{\alpha} - (p_{A1} - c_{A1}) (1 - s_{A1}) + (p_{A2} - c_{A2}) s_{A2} = 0$$

Similarly, the f.o.c. of profit maximization with respect to  $p_{A2}$  is:  $q_{A2} + p_{A2} \frac{\partial q_{A2}}{\partial p_{A2}} - c_{A2} \frac{\partial q_{A2}}{\partial p_{A2}} + p_{A1} \frac{\partial q_{A1}}{\partial p_{A2}} - c_{A1} \frac{\partial q_{A1}}{\partial p_{A2}} = 0$ . Following the same procedure as for the other condition, we obtain the following expression for the f.o.c:

$$\frac{1}{\alpha} + (p_{A1} - c_{A1}) s_{A1} + (p_{A2} - c_{A2}) (1 - s_{A2}) = 0$$

(b) Given this equation, write two equations for the equilibrium price-cost margins  $PCM_{A1} \equiv p_{A1} - c_{A1}$  and  $PCM_{A2} \equiv p_{A2} - c_{A2}$ , respectively. In these equations, the price cost margins are functions of  $s_{A1}$ ,  $s_{A2}$ , and the demand parameter  $\alpha$ . [Hint: Solve a system of two linear equations and two unknowns where the unknowns are the price cost margins and  $PCM_{A1}$  and  $PCM_{A2}$ ].

ANSWER.

(b) We have the following system of two linear equations with two unknowns, where the unknowns are  $PCM_{A1}$  and  $PCM_{A2}$ :

$$\begin{cases} PCM_{A1} (1 - s_{A1}) - PCM_{A2} s_{A2} &= \frac{1}{\alpha} \\ -PCM_{A1} s_{A1} + PCM_{A2} (1 - s_{A2}) &= \frac{1}{\alpha} \end{cases}$$

Solving this system of equations, we obtain:

$$PCM_{A1} = PCM_{A2} = \frac{1}{\alpha(1 - s_{A1} - s_{A2})}$$

**Q1.4. (10 points)** Suppose that products  $A1$  and  $A2$  were produced by two different firms that operate separately and maximize their respective profits. Answer the same Questions as in Q1.3(a) and Q1.3(B) but for these independent firms.

ANSWER.

(a) For any of the two firms  $j = A1$  or  $j = A2$ , the profit function is  $\pi_j = (p_j - c_j)q_j - FC(x_j)$ . The f.o.c. of profit maximization for firm  $j$  is:  $q_j + (p_j - c_j) \frac{\partial q_j}{\partial p_j}$ . This F.O.C. is:

$$q_j - \alpha q_j (p_j - c_j) (1 - s_j) = 0$$

(b) The f.o.c. for firm  $A1$  implies:

$$PCM_{A1} = \frac{1}{\alpha(1 - s_{A1})}$$

And similarly, the F.O.C. for firm  $A2$  implies:

$$PCM_{A2} = \frac{1}{\alpha(1 - s_{A2})}$$

**Q1.5. (10 points)** Compare the expressions for the equilibrium price cost margins in Questions Q1.3(b) and Q1.4(b).

**(a) Does the multi-product firm charge higher or lower price-cost margins than the single product firm?**

ANSWER.

(a) Since  $s_{A1} > 0$  and  $s_{A2} > 0$ , it is clear that  $1 - s_{A1} - s_{A2} < 1 - s_{A1}$  and  $1 - s_{A1} - s_{A2} < 1 - s_{A2}$ , such that:

$$PCM_{A1}^{Multiproduct} = \frac{1}{\alpha(1 - s_{A1} - s_{A2})} > \frac{1}{\alpha(1 - s_{A1})} = PCM_{A1}^{Single-product}$$

$$PCM_{A2}^{Multiproduct} = \frac{1}{\alpha(1 - s_{A1} - s_{A2})} > \frac{1}{\alpha(1 - s_{A2})} = PCM_{A2}^{Single-product}$$

Therefore, the multiproduct firm charges higher price-cost margins than the single-product firm.

**(b) Based on this result, explain in words the implications of multiproduct firms on prices, firms' profits, and consumer surplus.**

ANSWER.

(b) Prices. Given the same costs, a multi-product firm selling substitute products charges higher prices than the corresponding single-product firms. This is because the multi-product firm internalizes the substitution effects (or cannibalization effects) of the products it sells. Reducing the price  $p_{A1}$  can increase the demand and the profit from selling product  $A1$ , but it also reduces the demand of product  $A2$  and potentially also the profit from selling this product.

Profits. The profit of a multiproduct firm is higher than the sum of the corresponding single-product firms. It is clear that the multi-product firm can charge the same prices as the single product firm and therefore it can achieve at least the same profit as the sum of the profits of the single-product firms. But it can do better than that because it internalizes the substitution effects between products and maximizes the joint profit.

Consumer welfare. Since the multi-product firm charges higher prices, we have that quantities sold are lower, and consumer surplus is also smaller than with single-product firms.

**(c) Suppose that there are economies of scope in the production of the two goods. Explain in words how the presence of economies of scope affect your answer to question (b).**

ANSWER.

(c) The arguments in point (b) are based on the assumption that the multi-product firm has the same marginal costs for each good as the single-product firm. That is,  $c_j$  is the same regardless the firm is single-product or multi-product. However, the production of some goods may share the same inputs, and there may be economies of scope. Under economies of scope, the multi-product firm has lower marginal costs than the corresponding single-product firm. This reduction in marginal costs has a negative effect on prices. This effect can partially or even fully offset the positive effect on prices due to the internalization of demand substitution effects. As a result, with economies of scope, a multi-product firm can charge either higher or lower prices than single-product firms. It depends on the magnitude of the economies of scope in costs and the degree of substitution between products in the demand.

**Q1.6. (20 points) Suppose that firms A and B merge into a single firm that maximizes the total profits from the four products.**

**(a) Obtain the expression for the price cost margin of each of the four products.**

ANSWER.

(a) Following the same approach as in Q1.3 above, we obtain that the new multi-product firm charges the same price-cost margins for the four product, and these margins are:

$$PCM_{A1} = PCM_{A2} = PCM_{B1} = PCM_{B2} = \frac{1}{\alpha(1 - s_{A1} - s_{A2} - s_{B1} - s_{B2})}$$

**(b) Under the assumption of no economies of scope, explain in words the implications of this merge on prices, profits, and consumer surplus.**

ANSWER.

(b) Since  $s_{A1} > 0$ ,  $s_{A2} > 0$ ,  $s_{B1} > 0$ , and  $s_{B2} > 0$ , it is clear that the new firm after the merger has higher margins than the two separate firms. Since marginal costs are the same (no economies of scope), we have that prices are higher, profits increase, and consumer surplus declines.

**(c) Suppose that there are economies of scope. Explain in words the implications on the welfare effects of a merger.**

ANSWER.

(c) With economies of scope, the new firm after the merger has lower marginal costs and this can partially or completely offset the effects in point (b).

**QUESTION 2.** [80 points]. To answer the questions in this part of the problem set you need to use the dataset `verboven_cars.dta`. Use this dataset to implement the estimations describe below. Please, provide the STATA code that you use to obtain the results. For all the models that you estimate below, impose the following conditions:

- For market size (number of consumers), use  $\text{Population}/4$ , i.e., `pop/4`
- Use prices measured in euros (`eurpr`).
- For the product characteristics in the demand system, include the characteristics: `hp`, `li`, `wi`, `cy`, `le`, and `he`.
- Include also as explanatory variables the market characteristics:  $\log(\text{pop})$  and  $\log(\text{gdp})$ .
- In all the OLS estimations include fixed effects for market (`ma`), year (`ye`), and brand (`brd`).
- Include the price in logarithms, i.e.,  $\log(\text{eurpr})$ .
- Allow the coefficient for log-price to be different for different markets (countries). That is, include as explanatory variables the log price, but also the log price interacting (multiplying) each of the market (country) dummies except one country dummy (say the dummy for Germany) that you use as a benchmark.

**Q2.1. (10 points)**

ANSWER. The complete code is provided at the end of this document. Here I include some parts of the code. First, we read the dataset and construct some new variables.

```
use "C:\verboven_cars.dta", clear
gen logq = ln(qu)
gen logp = ln(eurpr)
gen logpop = ln(pop)
gen loggdp = ln(ngdp)
gen msize = pop/4
```

We need to construct the variable with the market shares ( $s_{jmt}$ ), the market share of the outside alternative ( $s_{0mt}$ ), and the log-odds ratio ( $\ln(s_{jmt}/s_{0mt})$ ).

```
//construct market share s_j
```

```

gen share = qu/msize
//construct outside good's market share s_0
egen sum_share = sum(share), by(ma ye)
gen share0 = 1 - sum_share
//generate log odd ratio
gen lsj_ls0 = ln(share/share0)

```

And we need to generate the dummy variables for country (market) and the product of each of these dummies with log-price.

```

//generate country dummies
tab ma, gen(countrydum_)
gen dumBel_logp = countrydum_1 * logp
gen dumFra_logp = countrydum_2 * logp
gen dumGer_logp = countrydum_3 * logp
gen dumIta_logp = countrydum_4 * logp
gen dumUK_logp = countrydum_5 * logp

```

**(a) Obtain the OLS-Fixed effects estimator of the Standard logit model. Interpret the results.**

ANSWER. The code is (including logp and omitting Germany \* logp)

```

// With logp and Germany as excluding dummy * logp
reghdfe lsj_ls0 logp dumBel_logp dumFra_logp dumIta_logp dumUK_logp hp li wi
cy le he logpop loggdp, vce(robust) a(ma ye brd)

```

This is the table with the estimation results.



```

. // With logp and Germany as excluding dummy * logp
. reghdfe lsj_ls0 logp dumBel_logp dumFra_logp dumIta_logp dumUK_logp hp li wi cy le he logpop loggdp
(converged in 8 iterations)

HDFE Linear regression                               Number of obs   =       11,549
Absorbing 3 HDFE groups                             F(      13,   11463) =       243.25
Statistics robust to heteroskedasticity              Prob > F           =        0.0000
                                                    R-squared          =        0.4099
                                                    Adj R-squared      =        0.4056
                                                    Within R-sq.       =        0.2310
                                                    Root MSE          =        1.1578

```

lsj_ls0	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
logp	-.6834595	.1202965	-5.68	0.000	-.9192613 -.4476577
dumBel_logp	-.3969749	.0573774	-6.92	0.000	-.5094444 -.2845054
dumFra_logp	-.7097122	.0636482	-11.15	0.000	-.8344736 -.5849507
dumIta_logp	-.8742421	.0857869	-10.19	0.000	-1.042399 -.7060851
dumUK_logp	-.5035495	.0713707	-7.06	0.000	-.6434483 -.3636508
hp	-.013324	.0017628	-7.56	0.000	-.0167794 -.0098685
li	-.0408816	.0136327	-3.00	0.003	-.067604 -.0141592
wi	.0648839	.0033021	19.65	0.000	.0584111 .0713567
cy	-.0007224	.0000852	-8.48	0.000	-.0008893 -.0005554
le	-.0001112	.0007775	-0.14	0.886	-.0016352 .0014128
he	-.0173409	.0029955	-5.79	0.000	-.0232125 -.0114693
logpop	-.7868595	.2654413	-2.96	0.003	-1.30717 -.2665493
loggdp	.8262299	.0877512	9.42	0.000	.6542226 .9982372

Absorbed degrees of freedom:

Absorbed FE	Num. Coefs. =	Categories -	Redundant
ma	5	5	0
ye	29	30	1
brd	39	40	1 ?

? = number of redundant parameters may be higher

Alternatively, but with identical results, we can implement this estimator without including logp and including all the country dummies interacted with logp. The code is:

```

. // Without logp and including all country dummies * logp
. reghdfe lsj_ls0 dumGer_logp dumBel_logp dumFra_logp dumIta_logp dumUK_logp hp li wi cy le he logpop
(converged in 8 iterations)

HDFE Linear regression                               Number of obs   =       11,549
Absorbing 3 HDFE groups                             F(      13,   11463) =       243.25
Statistics robust to heteroskedasticity              Prob > F           =        0.0000
                                                    R-squared          =        0.4099
                                                    Adj R-squared      =        0.4056
                                                    Within R-sq.       =        0.2310
                                                    Root MSE          =        1.1578

```

lsj_ls0	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
dumGer_logp	-.6834595	.1202965	-5.68	0.000	-.9192613 -.4476577
dumBel_logp	-1.080434	.1088015	-9.93	0.000	-1.293704 -.8671648
dumFra_logp	-1.393172	.1140136	-12.22	0.000	-1.616658 -1.169686
dumIta_logp	-1.557702	.1150547	-13.54	0.000	-1.783228 -1.332175
dumUK_logp	-1.187009	.1075541	-11.04	0.000	-1.397833 -.9761846
hp	-.013324	.0017628	-7.56	0.000	-.0167794 -.0098685
li	-.0408816	.0136327	-3.00	0.003	-.067604 -.0141592
wi	.0648839	.0033021	19.65	0.000	.0584111 .0713567
cy	-.0007224	.0000852	-8.48	0.000	-.0008893 -.0005554
le	-.0001112	.0007775	-0.14	0.886	-.0016352 .0014128
he	-.0173409	.0029955	-5.79	0.000	-.0232125 -.0114693
logpop	-.7868595	.2654413	-2.96	0.003	-1.30717 -.2665493
loggdp	.8262299	.0877512	9.42	0.000	.6542226 .9982372

Absorbed degrees of freedom:

Absorbed FE	Num. Coefs. =	Categories -	Redundant
ma	5	5	0
ye	29	30	1
brd	39	40	1 ?

? = number of redundant parameters may be higher

Interpretation:

- In the estimation without including  $\log p$  and including all the country dummies interacted with  $\log p$ , the parameter estimates are estimates of  $\alpha_{Germany}$ ,  $\alpha_{Belgium}$ ,  $\alpha_{France}$ ,  $\alpha_{Italy}$ , and  $\alpha_{UK}$ . In the estimation including  $\log p$  and omitting  $Germany * \log p$ , the parameter estimate associated to regressor  $\log p$  is  $\alpha_{Germany}$ , and the parameter estimates for the interactions of country dummies with  $\log p$  are  $(\alpha_{Belgium} - \alpha_{Germany})$ ,  $(\alpha_{France} - \alpha_{Germany})$ ,  $(\alpha_{Italy} - \alpha_{Germany})$ , and  $(\alpha_{UK} - \alpha_{Germany})$ .

- The parameters  $\alpha$  that measure price sensitivity are all significantly smaller than zero as we expect with a downward sloping demand curve. There are differences between countries in these parameters. Below we present a formal test of the null hypothesis that these parameters are the same. But it is already clear that the parameter  $\alpha$  for Germany is substantially smaller in absolute value than the corresponding parameters for other countries. Italy is the market with the largest price sensitivity.

- The parameter estimate for the characteristic "width" ( $w_i$ ) is positive and statistically significant. However, the parameter estimates of the other product characteristics are all negative and most of them significant. This result is not reasonable, especially for characteristics such as cylinder volume ( $cy$ ) and horsepower ( $hp$ ) which are measures of quality. There are several possible interpretation for this implausible estimates of these parameters. Let me include here two possible explanations. First, consumers may have heterogeneous taste for these characteristics. Some consumers like them and others dislike them. We are estimating just the average taste. Though this is possible, it is hard to believe that the average consumer valuation of cylinder volume ( $cy$ ) or horsepower ( $hp$ ) is negative. A second and more plausible explanation is that some of these observable characteristics are negatively correlated with the unobserved quality of the product,  $\xi_j$ . The estimated parameters are capturing both the direct (*ceteris paribus*) effect of these characteristics on demand, but also the indirect (non causal) effect because their correlation with the error term.

- The effect of  $\log gdp$  is positive on the demand of cars. Markets and time periods with more  $\log gdp$  have a larger demand for all the products. This positive income effect makes economic sense.

- Instead,  $\log pop$  has a negative effect on the demand for cars. This could be interpreted as a causal effect: markets with more population tend to have consumers with a lower taste for cars. But it can be also interpreted as a correction for our measure of market size,  $pop/4$ .

**(b) Test the null hypothesis that all countries have the same price coefficient.**

ANSWER. The implementation of the test depends on how we have implemented the estimation: including  $\log p$  and omitting one country dummy; or omitting  $\log p$  and including all the country dummies.

If we include `logp` and omit one `Germany*logp`, this is the code for the test.

```
// Test of null hypothesis same alphas across countries
test dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp = 0
```

And this is the result or output from STATA:

```
. test dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp = 0

( 1)  dumBel_logp - dumFra_logp = 0
( 2)  dumBel_logp - dumIta_logp = 0
( 3)  dumBel_logp - dumUK_logp = 0
( 4)  dumBel_logp = 0

      F( 4, 11463) =    37.08
      Prob > F =    0.0000
```

If we do not include `logp` and include all the country dummies interacted with `logp`, this is the code for the test.

```
// Test of null hypothesis same alphas across countries
test dumGer_logp = dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp
```

And this is the result or output from STATA:

```
. test dumGer_logp = dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp

( 1)  dumGer_logp - dumBel_logp = 0
( 2)  dumGer_logp - dumFra_logp = 0
( 3)  dumGer_logp - dumIta_logp = 0
( 4)  dumGer_logp - dumUK_logp = 0

      F( 4, 11463) =    37.08
      Prob > F =    0.0000
```

Interpretation. The two approaches give us exactly the same result for the F-test. This should be necessarily the case by construction. The p-value of this null hypothesis is practically zero. Therefore, under practically any significance level, we reject the null hypothesis that the  $\alpha$  parameters are the same across countries.

(c) Based on the estimated model, obtain the average price elasticity of demand for each country evaluated at the mean values of prices and market shares for that country.

ANSWER. The demand elasticity for product-country-year  $(j, m, t)$  is:

$$\eta_{jmt} = \frac{\partial s_{jmt}}{\partial p_{jmt}} \frac{p_{jmt}}{s_{jmt}}$$

Now, we take into account the form of  $\partial s_{jmt} / \partial p_{jmt}$  in our Logit model.

$$\frac{\partial s_{jmt}}{\partial p_{jmt}} = \frac{\partial s_{jmt}}{\partial \delta_{jmt}} \frac{\partial \delta_{jmt}}{\partial p_{jmt}}$$

In the Logit model,  $\frac{\partial s_{jmt}}{\partial \delta_{jmt}} = s_{jmt}(1 - s_{jmt})$ . And given that  $\delta_{jmt} = -\alpha_m \ln(p_{jmt}) + \dots$ , we have that  $\frac{\partial \delta_{jmt}}{\partial p_{jmt}} = \frac{-\alpha_m}{p_{jmt}}$ . Therefore,

$$\eta_{jmt} = \left[ s_{jmt}(1 - s_{jmt}) \frac{-\alpha_m}{p_{jmt}} \right] \frac{p_{jmt}}{s_{jmt}} = -\alpha_m (1 - s_{jmt})$$

As for the code, we start by creating a variable that contains the information about the countries' alpha parameters. This is the code.

```
// Creating variable with alpha parameters
gen alpha = .
replace alpha = _b[dumBel_logp] if ma==1
replace alpha = _b[dumFra_logp] if ma==2
replace alpha = _b[dumGer_logp] if ma==3
replace alpha = _b[dumIta_logp] if ma==4
replace alpha = _b[dumUK_logp] if ma==5
```

Variable **alpha** contains the estimated parameter  $\alpha_{Belgium}$  for all the observations that belong to Belgium, the estimated parameter  $\alpha_{France}$  for all the observations that belong to France, and so on for each country. Next, we create a variable that contains the values of the elasticities  $\eta_{jmt} = -\alpha_m (1 - s_{jmt})$  for every product-country-year observation in the data. This is the code.

```
// Creating variable with the elasticities
gen elasticity = -alpha * (1-share)
```

To obtain the mean value of this variable by country, we could do it in different ways. For instance,

```
// Mean elasticities by country
sum elasticity if ma==1
```

```

sum elasticity if ma==2
sum elasticity if ma==3
sum elasticity if ma==4
sum elasticity if ma==5

```

Or in a more compact form

```

// Mean elasticities by country
tab ma, sum(leasticity)

```

These are the results

```

. tab ma, sum(elasticity)

```

market (=second dimension of panel)	Summary of elasticity		
	Mean	Std. Dev.	Freq.
Belgium	-1.0787248	.00193667	2,673
France	-1.3908184	.00384684	2,265
Germany	-.68220874	.00203348	2,283
Italy	-1.5550288	.0050353	2,027
UK	-1.1853625	.00248949	2,301
Total	-1.1663937	.29085748	11,549

#### (d) Interpret the results

ANSWER. (d) Interpretation of the results. In principle, the differences in the mean demand elasticities across countries come from two sources: (1) differences between countries in the  $\alpha$  parameters; and (2) differences between countries in the mean market shares. However, mean market shares are very similar across countries. Therefore, mean demand elasticities are practically identical to the  $\alpha$  parameters.

**Q2.2. (10 points) Consider the equilibrium condition (first order conditions of profit maximization) under the assumption that each product is produced by only one firm.**

**(a) Write the equation for this equilibrium condition. Write this equilibrium condition as an equation for the Lerner Index,  $\frac{p_j - MC_j}{p_j}$ .**

ANSWER. The profit function of firm  $j$  is  $\pi_j = p_j q_j - C_j(q_j)$ . The first order condition with respect to price is:

$$q_j + p_j \frac{\partial q_j}{\partial p_j} - MC_j \frac{\partial q_j}{\partial p_j} = 0$$

Solving for  $p_j - MC_j$ , we have that:

$$p_j - MC_j = \frac{-q_j}{\partial q_j / \partial p_j}$$

And the Lerner index is:

$$\frac{p_j - MC_j}{p_j} = \frac{-q_j}{p_j [\partial q_j / \partial p_j]}$$

Now, we take into account the form of  $\partial q_j / \partial p_j$  in our Logit model.

$$\frac{\partial q_j}{\partial p_j} = H \frac{\partial s_j}{\partial p_j} = H \frac{\partial s_j}{\partial \delta_j} \frac{\partial \delta_j}{\partial p_j}$$

In the Logit model,  $\frac{\partial s_j}{\partial \delta_j} = s_j(1 - s_j)$ . And given that  $\delta_j = -\alpha \ln(p_j) + \dots$ , we have that

$$\frac{\partial \delta_j}{\partial p_j} = \frac{-\alpha}{p_j}. \text{ Therefore,}$$

$$\frac{\partial q_j}{\partial p_j} = q_j (1 - s_j) \frac{-\alpha}{p_j}$$

And plugging this expression in the Lerner Index, we have:

$$\frac{p_j - MC_j}{p_j} = \frac{-q_j}{p_j [q_j (1 - s_j) \frac{-\alpha}{p_j}]} = \frac{1}{\alpha(1 - s_j)}$$

**(b) Using the previous equation in Q2.2(a) and the estimated demand in Q2.1, calculate the Lerner index for every car-market-year observation in the data.**

ANSWER. Note that, according to the estimated models, different countries have different value for the parameter  $\alpha$ . Furthermore, every product-market-year observation  $(j, m, t)$  has a different market share,  $s_{jmt}$ . Therefore, the Lerner index variable is:

$$\frac{p_{jmt} - MC_{jmt}}{p_{jmt}} = \frac{1}{\alpha_m (1 - s_{jmt})}$$

Note that, in this model, the Lerner index is exactly equal to the inverse of absolute value of the demand elasticity:

$$\frac{p_{jmt} - MC_{jmt}}{p_{jmt}} = \frac{1}{|\eta_{jmt}|}$$

This is the code to obtain the Lerner index variable.

```
// Lerner indexes
gen lerner = 1/abs(elasticity)
```

(c) Report the mean values of the Lerner Index for each of the counties/markets. Comment the results.

ANSWER. Again there are different way of computing the mean values of the Lerner index by country. We can obtain them as follows:

```
// Mean lerner indexes by country
// Note that Belgium (ma==1), France (ma==2),
// Germany (ma==3), Italy (ma==4), and UK (ma==5)
sum lerner if ma==1
sum lerner if ma==2
sum lerner if ma==3
sum lerner if ma==4
sum lerner if ma==5
```

Or in a more compact form:

```
// Mean lerner indexes by country
tab ma, sum(lerner)
```

These are the results:  

```
. tab ma, sum(lerner)
```

market (=second dimension of panel)	Summary of lerner		Freq.
	Mean	Std. Dev.	
Belgium	.92702347	.00167203	2,673
France	.71900666	.00200461	2,265
Germany	1.4658401	.00441103	2,283
Italy	.64308173	.00211327	2,027
UK	.84362754	.00178202	2,301
Total	.9262889	.28535342	11,549

.

The estimated demand model implies Lerner indexes that are very high, especially for Germany where the index implies a markup of 146%. This is because the elasticities are small, especially in Germany.

**Q2.3. (20 points)**

**(a) Using the equilibrium condition and the estimated demand, obtain an estimate of the marginal cost for every car-market-year observation in the data.**

ANSWER. The equilibrium condition of the model implies that:

$$p_{jmt} - MC_{jmt} = \frac{p_{jmt}}{\alpha_m (1 - s_{jmt})}$$

And solving for the marginal cost:

$$MC_{jmt} = p_{jmt} - \frac{p_{jmt}}{\alpha_m (1 - s_{jmt})}$$

By looking at this expression, we can already see that the small estimate of the parameter  $\alpha_{Germany}$  (smaller than one) implies that the estimated marginal costs for Germany are negative, i.e., because with  $\alpha < 1$ ,  $p/\alpha(1 - s) > p$ . Of course, this is a very implausible prediction. There are two main aspects where the model could be failing and generating this implausible prediction: (1) ignoring that firms are multi-product; and (2) the logit model is very restrictive. A nested logit that accounts for the multiproduct nature of these firms does not generate these negative estimates for the marginal costs.

This is the code for calculating the marginal cost.

```
// Estimating the values of Marginal Costs
gen mc = eurpr - eurpr/(alpha * (1-share))
```

**(b) Run an OLS-Fixed effects regression where the dependent variable is the estimated value of the marginal cost, and the explanatory variables (regressors) are the product characteristics hp, li, wi, cy, le, and he. Interpret the results.**

ANSWER. This is the code for the estimation of the regression for the marginal cost function.

```
// Estimating marginal cost function
reghdfe mc hp li wi cy le he, vce(robust) a(ma ye brd)
```

And this is the output from this regression:



```
. // Estimating marginal cost function
. reghdfe mc hp li wi cy le he, vce(robust) a(ma ye brd)
(converged in 8 iterations)
```

HDFE Linear regression	Number of obs	=	11,549
Absorbing 3 HDFE groups	F( 6, 11470)	=	27.08
Statistics robust to heteroskedasticity	Prob > F	=	0.0000
	R-squared	=	0.7028
	Adj R-squared	=	0.7008
	Within R-sq.	=	0.0201
	Root MSE	=	1560.4886

mc	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
hp	15.13636	3.143366	4.82	0.000	8.974829	21.2979
li	6.80395	19.81002	0.34	0.731	-32.02708	45.63498
wi	1.46403	3.918408	0.37	0.709	-6.216719	9.144779
cy	-.4246081	.110773	-3.83	0.000	-.641742	-.2074742
le	2.03699	1.020322	2.00	0.046	.0369856	4.036995
he	2.574545	3.428862	0.75	0.453	-4.14661	9.2957

Interpretation. The product characteristic with the strongest effect on the marginal cost is horse power (**hp**). Since marginal cost (the dependent variable) is measured in Euros and horse power is measured in kW, we have that an increase in 1kW of horse power implies a 15 euros increase in the marginal cost of a car. The empirical distribution of horse power (`sum hp, detail`) has 10% percentile of 31 kW, a median of 55 kW, and a 90% percentile of 90 kW. Therefore, ceteris paribus, the difference in marginal cost between a car at the 90-percentile and a car at the 10-percentile of horse power is equal to  $15 \times (90 - 31) = 885$  euros. Length (**le**) has also a significant and positive effect on marginal cost. Since length (**le**) is measured in centimeters, we have that a 1 meter increase in length (which is approximately the difference between the 90-percentile and the 10-percentile in the empirical distribution of length) implies an increase in marginal cost of 203 euros. The characteristic cylinder (**cy**) has a negative effect on marginal cost, which is not a plausible prediction. A possible explanation is that this characteristic is negatively correlated with the error term, i.e., with omitted / unobserved product characteristics affecting marginal cost. The other characteristics have a positive effect but not statistically significant to zero.

We can obtain the estimates of country fixed effects, year fixed effects, and brand fixed effects, by using the option `savefe` in the estimation command (`reghdfe mc hp li wi cy le he, vce(robust) a(ma ye brd, savefe)`). Then, we can tabulate the values of these

variables:

```
tab ma, sum(maFE)
tab ye, sum(yeFE)
tab brd, sum(brdFE)
```

This shows that there is substantial heterogeneity in marginal costs (keeping constant the observable characteristics) across countries, and especially over brands. There is also a clear increasing time trend in the evolution of marginal costs over time.

**Q2.4. (40 points)** Consider the equilibrium condition (first order conditions of profit maximization) taking into account that each firm is a multi-product firm.

(a) Consider one firm, say 1, that sells products  $1, 2, \dots, J_1$ . Given our demand system here – that is different to the one in Question 1 – obtain the first order conditions of profit maximization for this firm. Use these conditions to obtain the following expression for the equilibrium value of the Lerner Index of a product  $j$ :

$$L_j \equiv \frac{p_j - MC_j}{p_j} = \frac{1}{\alpha} \left[ 1 + \frac{1}{p_j} \frac{\sum_{k=1}^{J_1} s_j p_j}{1 - \sum_{k=1}^{J_1} s_j} \right]$$

[Hint: First, use f.o.c. for product  $j$  to obtain an equation that relates  $L_j$  with the variable  $\Sigma \equiv \sum_{k=1}^{J_1} L_k s_k p_k$ . Second, aggregate this equation over products and solve for  $\Sigma$  to obtain its equilibrium values. Finally, plug in this value of  $\Sigma$  in the f.o.c. for product  $j$ ]

ANSWER. The profit function of firm 1 is  $\pi_1 = \sum_{j=1}^{J_1} p_j q_j - C_j(q_j)$ . The first order condition with respect to price  $j$  is:

$$q_j + \sum_{k=1}^{J_1} (p_k - MC_k) \frac{\partial q_k}{\partial p_j} = 0$$

In our demand model:

$$\frac{\partial q_j}{\partial p_j} = H \frac{\partial s_j}{\partial \delta_j} \frac{\partial \delta_j}{\partial p_j} \quad \text{and} \quad \frac{\partial q_k}{\partial p_j} = H \frac{\partial s_k}{\partial \delta_j} \frac{\partial \delta_j}{\partial p_j}$$

In the Logit model,  $\frac{\partial s_j}{\partial \delta_j} = s_j(1 - s_j)$  and  $\frac{\partial s_k}{\partial \delta_j} = -s_j s_k$ . And given that  $\delta_j = -\alpha \ln(p_j) + \dots$ , we have that  $\frac{\partial \delta_j}{\partial p_j} = \frac{-\alpha}{p_j}$ . Therefore,

$$\frac{\partial q_j}{\partial p_j} = q_j (1 - s_j) \frac{-\alpha}{p_j}$$

and

$$\frac{\partial q_k}{\partial p_j} = q_j s_k \frac{\alpha}{p_j}$$

Solving the expression in the F.O.C.  $q_j + \sum_{k=1}^{J_1} (p_k - MC_k) \frac{\partial q_k}{\partial p_j} = 0$ , we have that:

$$q_j - \alpha (p_j - MC_j) q_j (1 - s_j) \frac{1}{p_j} + \alpha \sum_{k \neq j} (p_k - MC_k) q_j s_k \frac{1}{p_j} = 0$$

This equation implies that:

$$L_j = \frac{1}{\alpha} + \frac{1}{p_j} \Sigma$$

where  $\Sigma \equiv \sum_{k=1}^{J_1} L_k s_k p_k$ . Multiplying LHS and RHS of this equation times  $s_j p_j$ , and then summing this equation over the  $J_1$  products. we get:

$$\Sigma = \frac{1}{\alpha} \left( \sum_{k=1}^{J_1} s_j p_j \right) + \Sigma \left( \sum_{k=1}^{J_1} s_j \right)$$

Solving for  $\Sigma$ , we get:

$$\Sigma = \frac{1}{\alpha} \frac{\sum_{k=1}^{J_1} s_j p_j}{1 - \sum_{k=1}^{J_1} s_j}$$

Finally, plugging this expression into equation  $L_j = \frac{1}{\alpha} + \frac{1}{p_j} \Sigma$ , we get:

$$L_j = \frac{1}{\alpha} \left[ 1 + \frac{1}{p_j} \frac{\sum_{k=1}^{J_1} s_j p_j}{1 - \sum_{k=1}^{J_1} s_j} \right]$$

**(b) Using this equilibrium condition for multi-product firms, the estimated demand, data on prices and market shares, and the actual ownership structure of products by the different firms, calculate the lerner-indexes, and the marginal costs for every car-market-year observation in the data.**

ANSWER. We construct the Lerner Index for a multiproduct firm by using the formula in Question 2.4(a) above. The variable in the dataset that represents the firm index is `fmr`.

We first generate the aggregate market share at the firm level:  $\sum_{k=1}^{J_1} s_j$

```
// Generating aggregate firm-level share
egen share_firm = sum(share), by(frm ma ye)
```

Second, we generate the aggregate firm-level weighted price:  $\sum_{k=1}^{J_1} s_j p_j$

```
// Generating aggregate firm-level weighted price
```

```
egen price_firm = sum(share * eurpr), by(firm ma ye)
```

Second, we generate the Lerner index.

```
// Generating Lerner Index for Multiproduct firm
gen lerner_multi = (1/alpha) * (1 + (1/eurpr) * price_firm/(1-share_firm))
```

To obtain the marginal cost implied by this Lerner index, we note that – by definition –  $L_j = (p_j - MC_j)/p_j$ , and this implies that  $MC_j = p_j(1 - L_j)$ .

```
// Generating Marginal Cost for Multiproduct firm
gen mc_multi = eurpr * (1 - lerner_multi)
```

**(c) Compare the estimated Lerner indexes under the assumption of single-product and under the assumption of multi-product. Present an xy scatterplot of these two variables. Comment the results.**

ANSWER. The following is the STATA code to generate the scatterplot for all the countries, and separated by country. We also include the 45-degree line to illustrate how the Lerner Index for a multiproduct firm is always larger than the Lerner index of single-product firm (for this demand system where all the products sold by a firm are substitutes with each other).

```
// All countries
scatter lerner_multi lerner_single
// Belgium
scatter lerner_multi lerner_single lerner_single if ma==1
// France
scatter lerner_multi lerner_single lerner_single if ma==2
// Germany
scatter lerner_multi lerner_single lerner_single if ma==3
// Italy
scatter lerner_multi lerner_single lerner_single if ma==4
// UK
scatter lerner_multi lerner_single lerner_single if ma==5
```

All countries

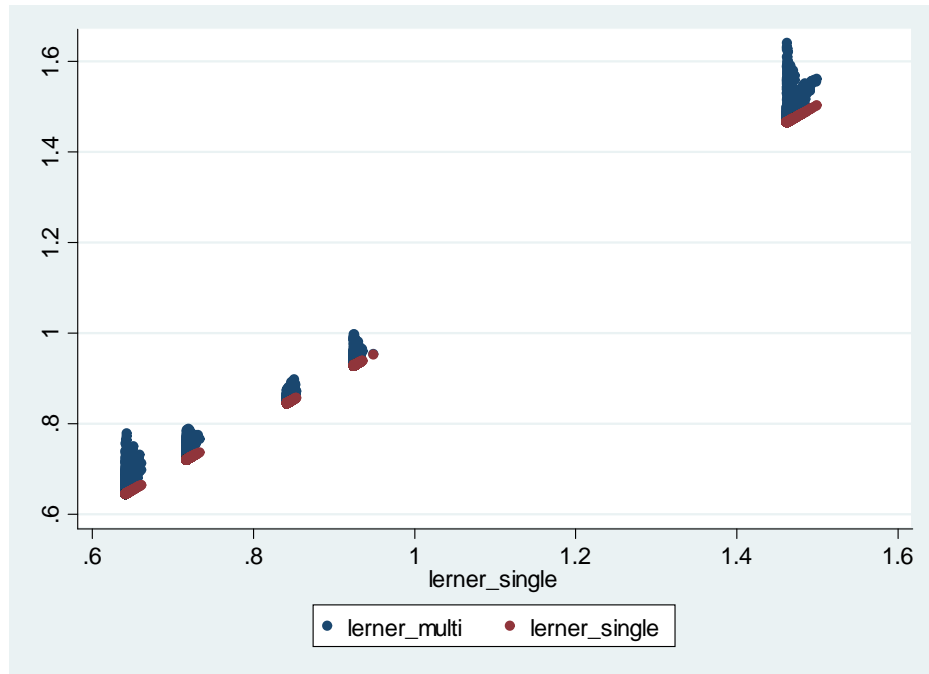


Figure for Belgium:

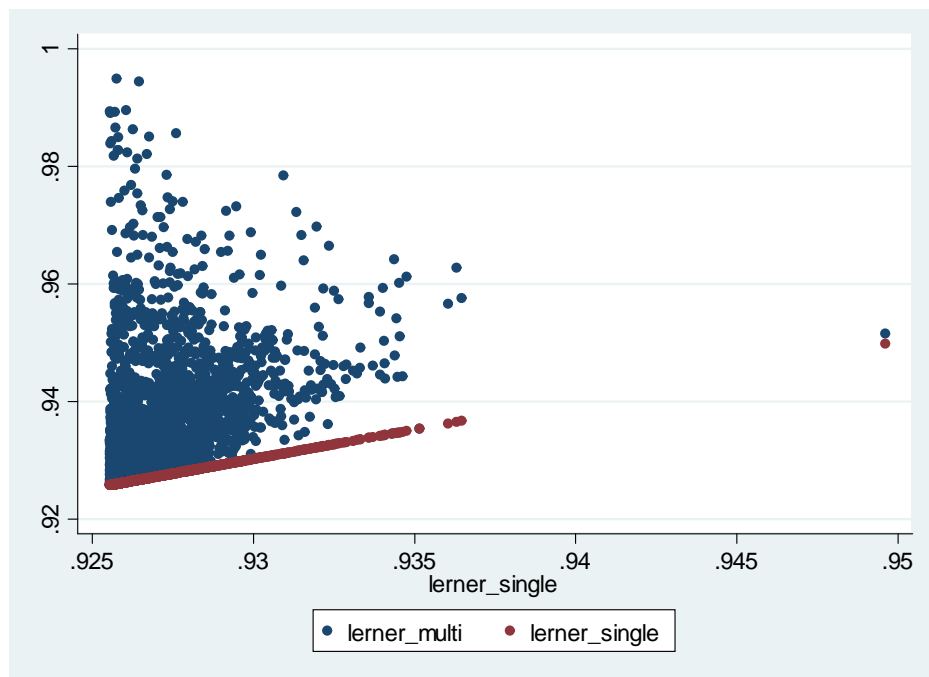


Figure for France:

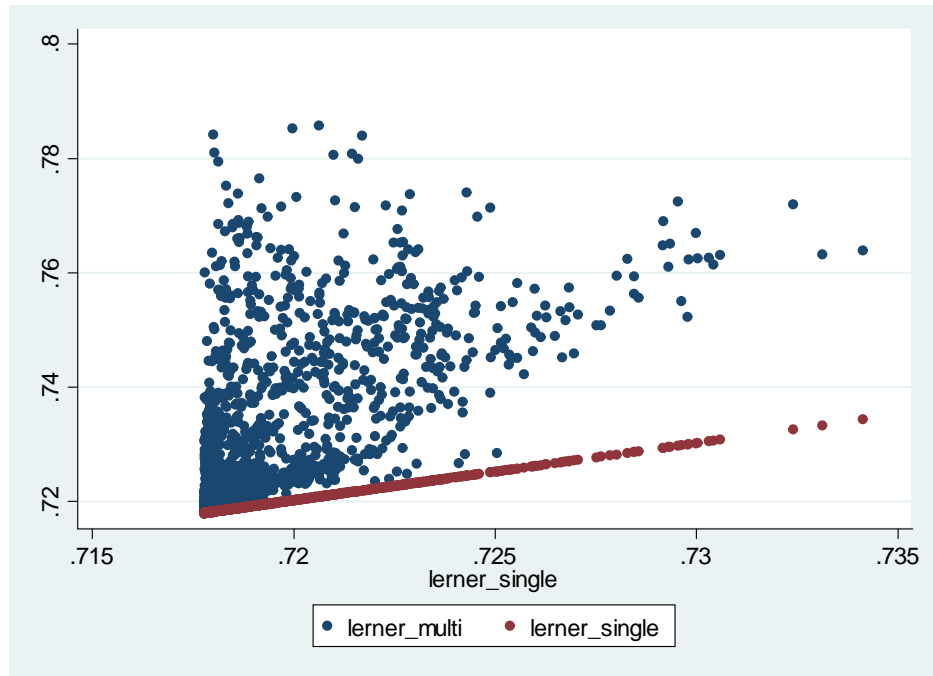


Figure for Germany:

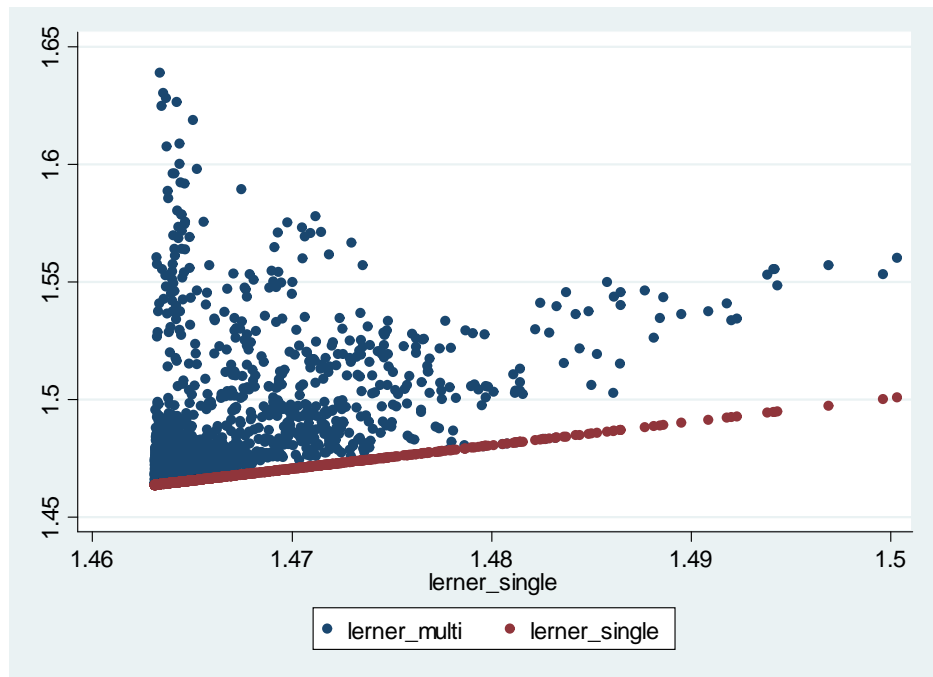


Figure for Italy:

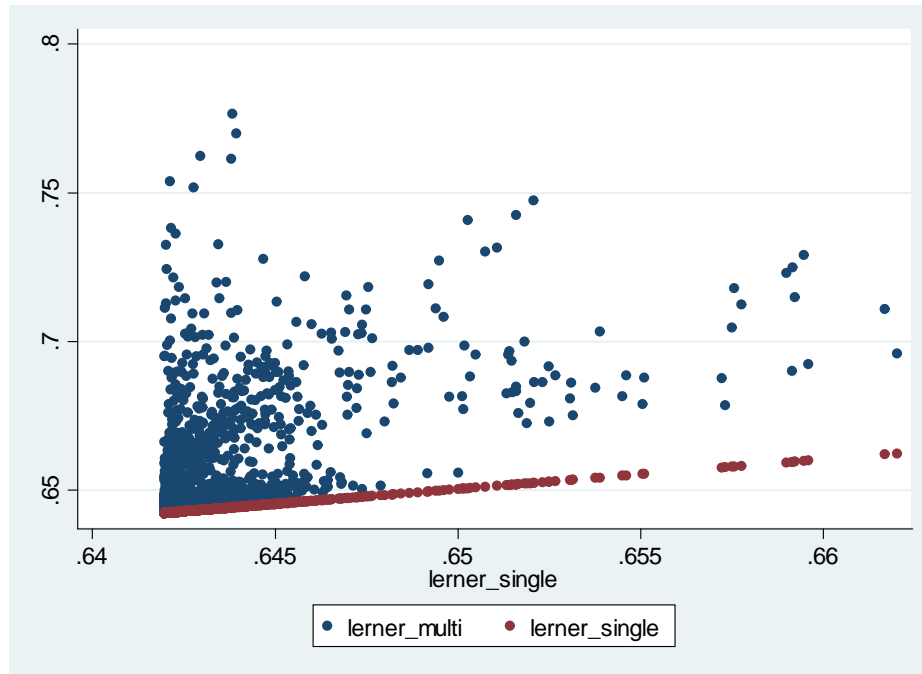
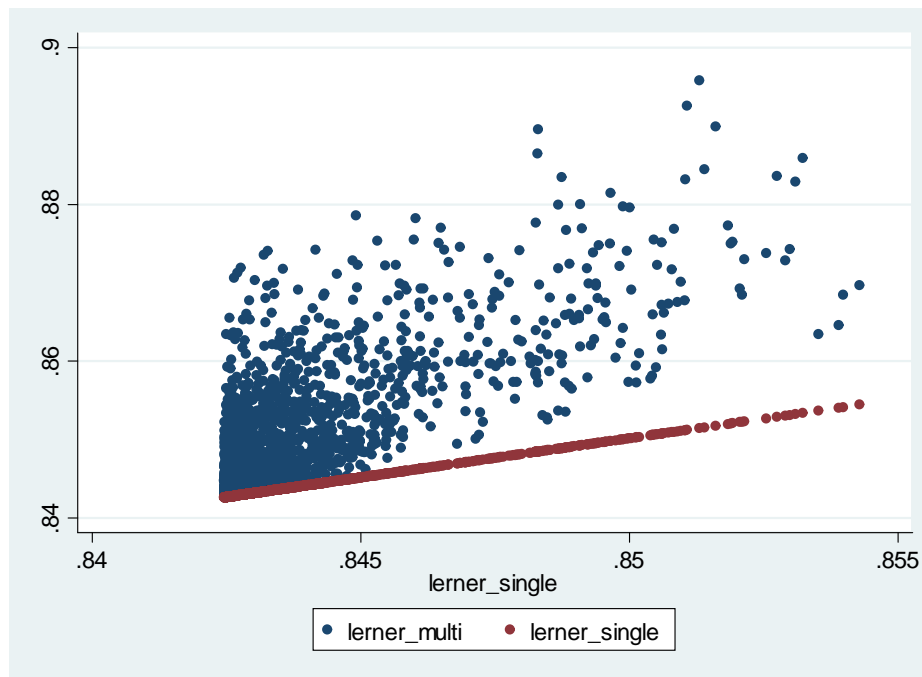


Figure for UK:



(d) Compare the estimated Marginal Costs under the assumption of single-product and under the assumption of multi-product. Present an xy scatterplot of these two variables. Comment the results.

ANSWER. The following is the STATA code to generate the scatterplot for all the countries, and separated by country. We also include the 45-degree line to illustrate how – by construction – the marginal cost estimated for a multiproduct firm is necessarily smaller than for a single-product firm. This is just a necessary implication of the differences between the Lerner indexes in question 2.4(c) above.

```
// All countries
scatter mc_multi mc_single
// Belgium
scatter mc_multi mc_single mc_single if ma==1
// France
scatter mc_multi mc_single mc_single if ma==2
// Germany
scatter mc_multi mc_single mc_single if ma==3
// Italy
scatter mc_multi mc_single mc_single if ma==4
// UK
scatter mc_multi mc_single mc_single if ma==5
```

All countries

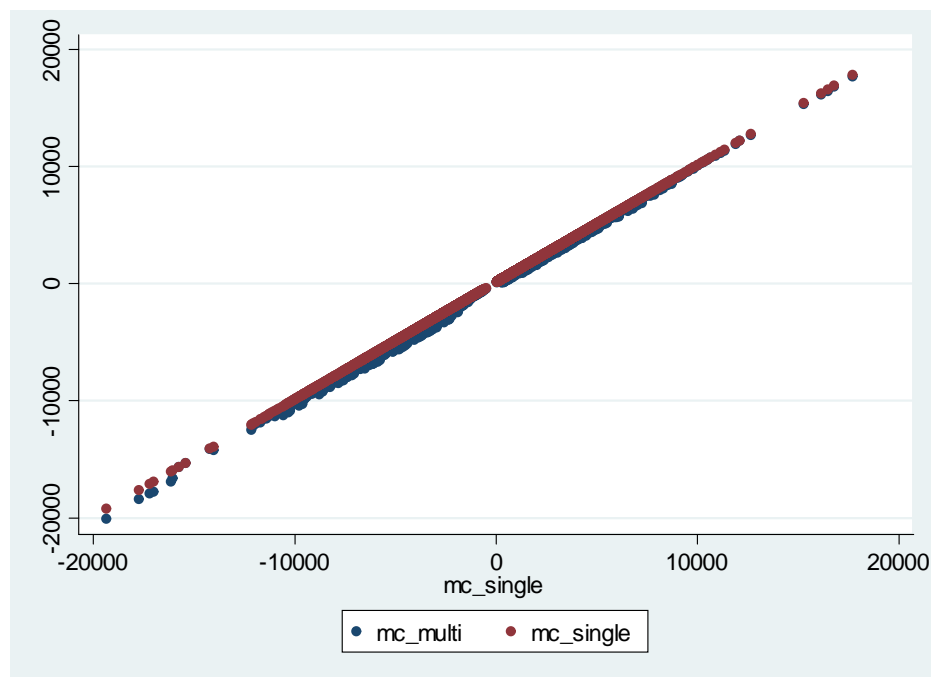


Figure for Belgium:



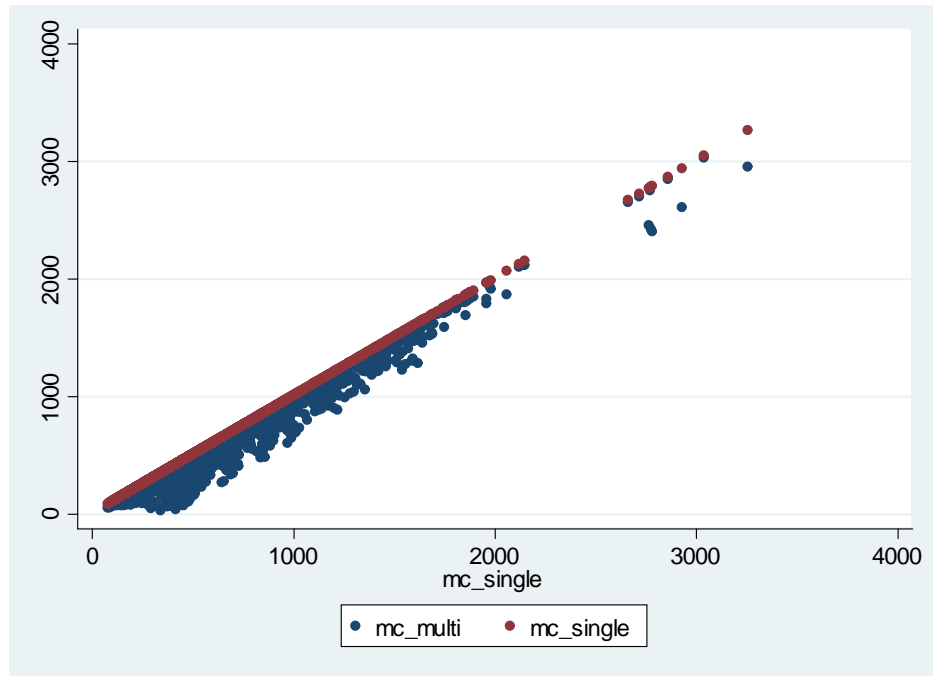


Figure for France:

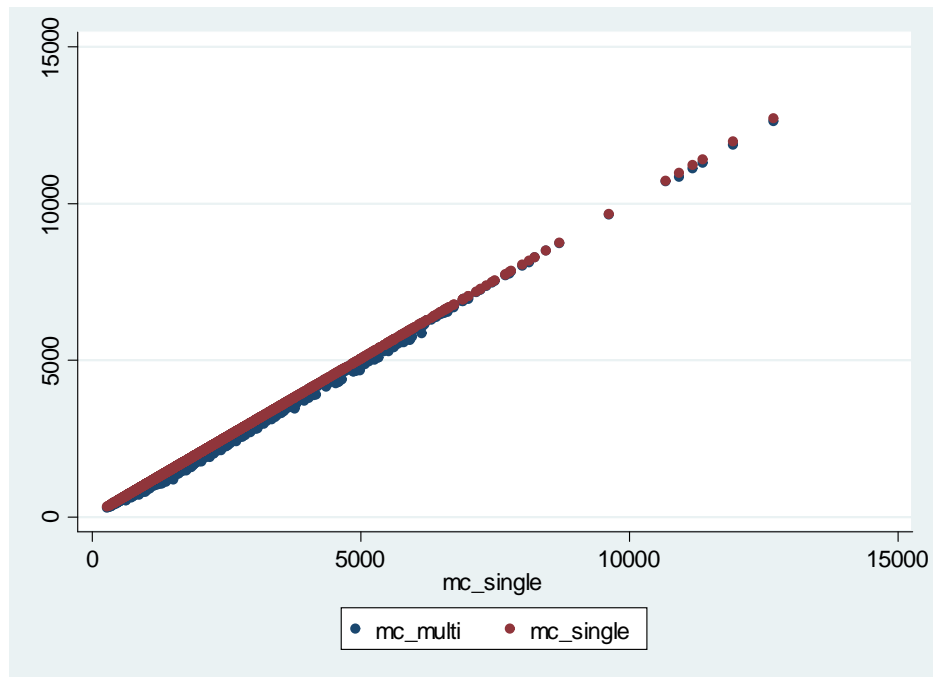


Figure for Germany:

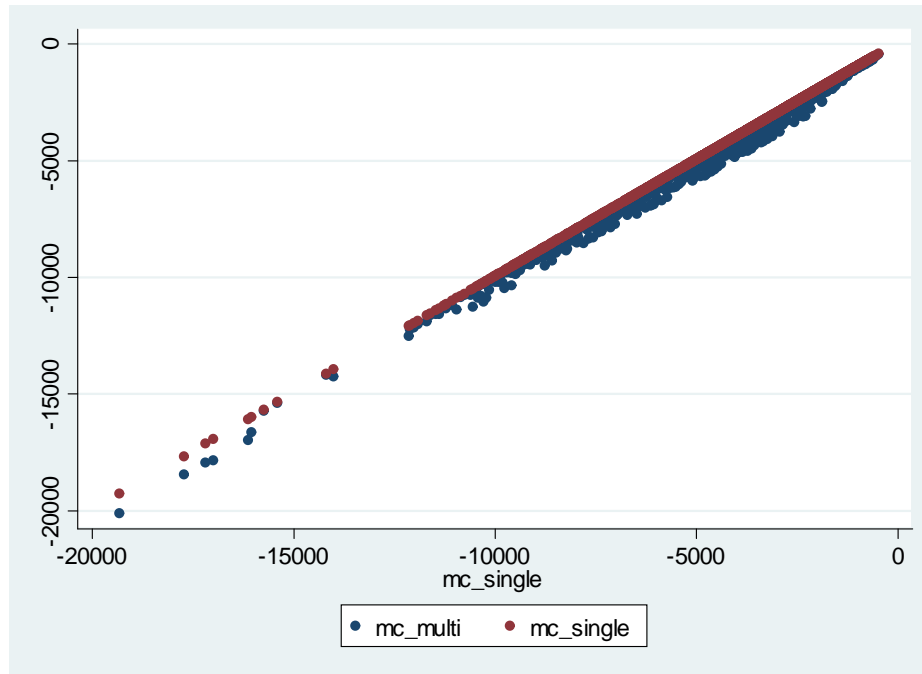


Figure for Italy:

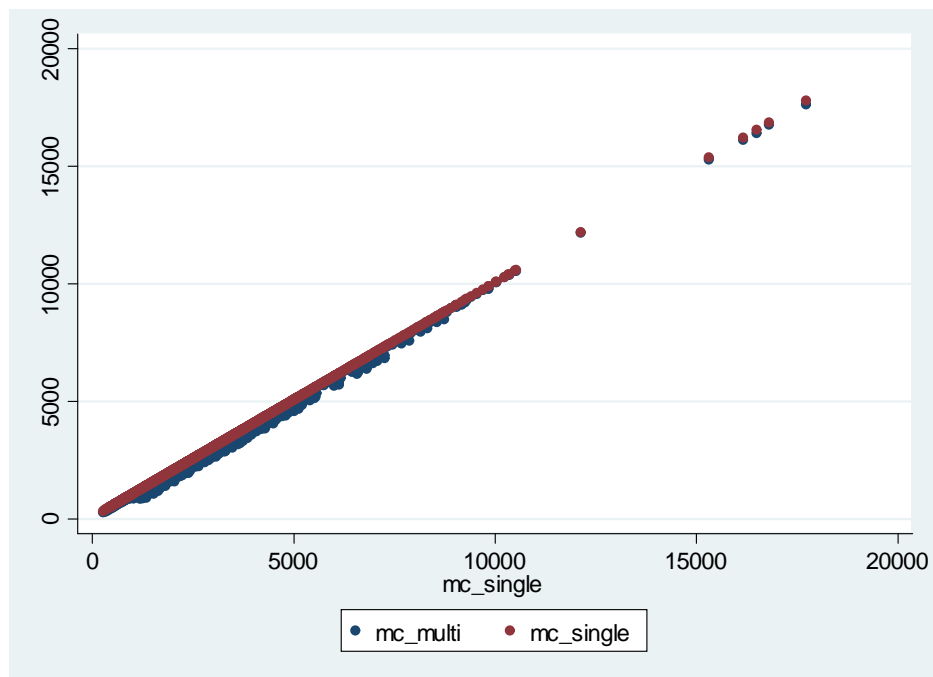
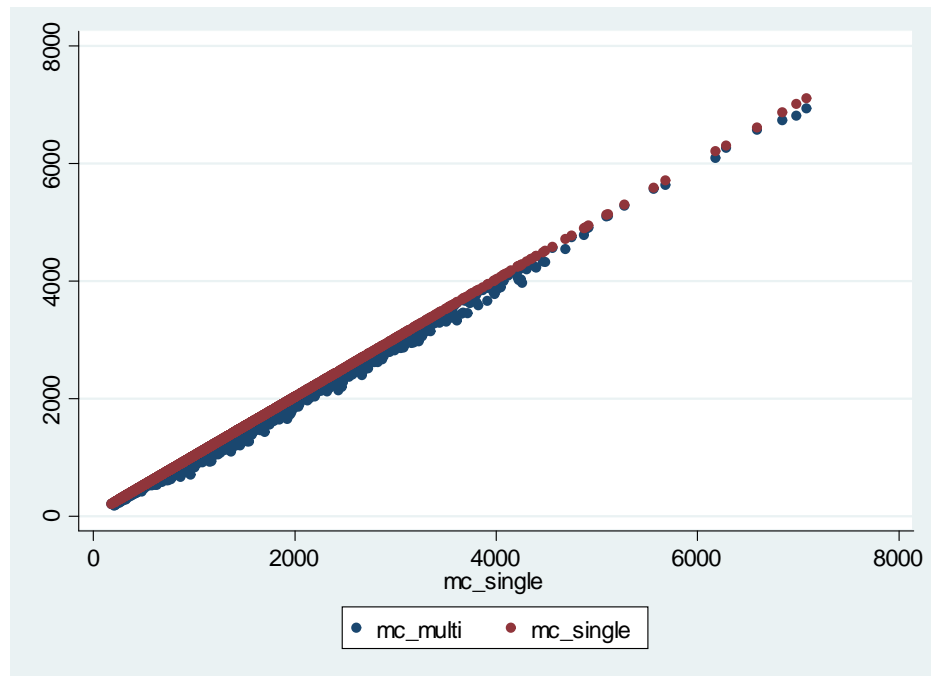


Figure for UK:



## STATA CODE

```
// -----  
// ECO310 - EMPIRICAL INDUSTRIAL ORGANIZATION  
//  
// PROBLEM SET 2  
//   using Verboven's data on automobiles in Europe  
// Due of December 5th, 2019  
//  
// by Victor Aguirregabiria  
//  
//-----  
//-----  
//   1. Reading dataset & generating new variables  
//-----  
use "C:\Users\Victor\Dropbox\documents\COURSES\ECO310\2018\PROBLEM_SETS\problem_set_01.dta"  
clear  
gen logq = ln(qu)  
gen logp = ln(eurpr)  
gen logpop = ln(pop)  
gen loggdp = ln(ngdp)  
gen msize = pop/4  
//construct market share s_j  
gen share = qu/msize  
//construct outside good's market share s_0  
egen sum_share = sum(share), by(ma ye)  
gen share0 = 1 - sum_share  
//generate log odd ratio  
gen lsj_ls0 = ln(share/share0)  
//generate country dummies  
tab ma, gen(countrydum_)  
// countrydum_1 ma==Belgium  
// countrydum_2 ma==France  
// countrydum_3 ma==Germany  
// countrydum_4 ma==Italy  
// countrydum_5 ma==UK
```

```

gen dumBel_logp = countrydum_1 * logp
gen dumFra_logp = countrydum_2 * logp
gen dumGer_logp = countrydum_3 * logp
gen dumIta_logp = countrydum_4 * logp
gen dumUK_logp = countrydum_5 * logp
//-----
//    3. Estimation of the Logit model and test of null hypothesis
//    Questions 2.1(a) and 2.1(b)
//-----
// With logp and Germany as excluding dummy * logp
reghdfe lsj_ls0 logp dumBel_logp dumFra_logp dumIta_logp dumUK_logp hp li wi
cy le he logpop loggdp, vce(robust) a(ma ye brd)
// Test of null hypothesis same alphas across countries
test dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp = 0
// Without logp and including all country dummies* logp
reghdfe lsj_ls0 dumGer_logp dumBel_logp dumFra_logp dumIta_logp dumUK_logp
hp li wi cy le he logpop loggdp, vce(robust) a(ma ye brd)
// Test of null hypothesis same alphas across countries
test dumGer_logp = dumBel_logp = dumFra_logp = dumIta_logp = dumUK_logp
//-----
//    4. Calculating demand elasticities
//    Elasticity =  $-\alpha * (1 - \text{share})$ 
//    Question 2.1(c)
//-----
tab ma, sum(ma)
// Creating variable with alpha parameters
gen alpha = .
replace alpha =  $-\_b[\text{dumBel\_logp}]$  if ma==1
replace alpha =  $-\_b[\text{dumFra\_logp}]$  if ma==2
replace alpha =  $-\_b[\text{dumGer\_logp}]$  if ma==3
replace alpha =  $-\_b[\text{dumIta\_logp}]$  if ma==4
replace alpha =  $-\_b[\text{dumUK\_logp}]$  if ma==5

// Creating variable with the elasticities
gen elasticity =  $-\alpha * (1 - \text{share})$ 
// Mean elasticities by country

```

```

sum elasticity if ma==1
sum elasticity if ma==2
sum elasticity if ma==3
sum elasticity if ma==4
sum elasticity if ma==5
// ... or in a more compact form
tab ma, sum(elasticity)
//-----
//    4. Lerner indexes and their means by country and car manufacturer
//        Question 2.2(b), 2.2(c), and 2.2(d)
//-----
gen lerner_single = 1/abs(elasticity)
// Mean lerner indexes by country
sum lerner_single if ma==1
sum lerner_single if ma==2
sum lerner_single if ma==3
sum lerner_single if ma==4
sum lerner_single if ma==5
// ... or in a more compact form
tab ma, sum(lerner_single)
// Aggregate sales by car manufacturer
egen totq_firm = sum(qu), by(frm)
// Top 5 manufacturers
gsort - totq_firm
list frm totq_firm if frm~=frm[_n-1]
// Mean lerner indexes for the top-5 firms
// Note that Fiat (frm==4), VW (frm==26),
// Peugeot (frm==16), Renault (frm==18), and Ford (frm==5)
tab frm, sum(frm)

tab frm if (frm==4) | (frm==26) | (frm==16) | (frm==18) | (frm==5), sum(lerner_single)
//-----
//    5. Estimating Marginal Costs and MArginal Cost function
//        Question 2.3(a) and 2.3(b)
//-----
// Estimating the values of marginal costs

```

```

gen mc_single = eurpr - eurpr/(alpha * (1-share))
// Estimating marginal cost function
reghdfe mc_single hp li wi cy le he, vce(robust) a(ma ye brd)
// Estimating marginal cost function
reghdfe mc_single hp li wi cy le he, vce(robust) a(ma ye brd, savefe)
// Estimating marginal cost function without Germany
reghdfe mc_single hp li wi cy le he if ma~=3, vce(robust) a(ma ye brd, savefe)

rename __hdf1__ maFE
rename __hdf2__ yeFE
rename __hdf3__ brdFE
tab ma, sum(maFE)
tab ye, sum(yeFE)
tab brd, sum(brdFE)
//-----
//    6. Calculating Lerner Indexes and Marginal Costs
//    for Multiproducts Firms
//        Question 2.4(b)
// "Calculate the lerner-indexs, and the marginal costs
//    for every car-market-year observation in the data."
//-----
// Variable that represents a firm is frm
// Generating aggregate firm-level share
egen share_firm = sum(share), by(frm ma ye)
// Generating aggregate firm-level weighted price
egen price_firm = sum(share * eurpr), by(frm ma ye)
// Generating Lerner Index for Multiproduct firm
gen lerner_multi = (1/alpha) * (1 + (1/eurpr) * price_firm/(1-share_firm))
// Generating Marginal Cost for Multiproduct firm
gen mc_multi = eurpr * (1 - lerner_multi)
//-----
//    7. Compare Lerner Indexes under single-product and multiproduct
//        Scatterplot
//        Question 2.4(c)
//-----
// All countries

```

```

scatter lerner_multi lerner_single
// Belgium
scatter lerner_multi lerner_single lerner_single if ma==1

// France
scatter lerner_multi lerner_single lerner_single if ma==2
// Germany
scatter lerner_multi lerner_single lerner_single if ma==3
// Italy
scatter lerner_multi lerner_single lerner_single if ma==4
// UK
scatter lerner_multi lerner_single lerner_single if ma==5
//-----
//      8.  Compare Marginal Costs under single-product and multiproduct
//          Scatterplot
//          Question 2.4(d)
//-----
// All countries
scatter mc_multi mc_single
// Belgium
scatter mc_multi mc_single mc_single if ma==1

// France
scatter mc_multi mc_single mc_single if ma==2
// Germany
scatter mc_multi mc_single mc_single if ma==3
// Italy
scatter mc_multi mc_single mc_single if ma==4
// UK
scatter mc_multi mc_single mc_single if ma==5

```