

Empirical Industrial Organization (ECO 310)

Winter 2021. Victor Aguirregabiria

Solution to Problem Set #2

Due on Friday, March 26, 2021 [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 Friday, March 26, 2021 [before 11:59pm].
6. Problem sets should be written individually.

The total number of marks is 200.

QUESTION 1. [80 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 ε_1 and ε_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 four products and average utilities $\delta_j = \beta x_j - \alpha p_j$ implies that the market share of product $A1$ is:

$$s_{A1} = \frac{\exp \{\delta_{A1}\}}{1 + \exp \{\delta_{A1}\} + \exp \{\delta_{A2}\} + \exp \{\delta_{B1}\} + \exp \{\delta_{B2}\}}$$

Q1.2. (15 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.

(a) By the chain rule, $\frac{\partial s_j}{\partial p_j} = \frac{\partial s_j}{\partial \delta_j} \frac{\partial \delta_j}{\partial p_j}$ where $\delta_j = \beta x_j - \alpha p_j$. Note that $\frac{\partial \delta_j}{\partial p_j} = -\alpha$ such that $\frac{\partial s_j}{\partial p_j} = -\alpha \frac{\partial s_j}{\partial \delta_j}$. We need to obtain $\frac{\partial s_j}{\partial \delta_j}$. To obtain this (partial) derivative, we look at s_j as a function only of δ_j and we can treat all the δ 's different to δ_j as constants. Therefore, we can write $s_j = \frac{\exp \{\delta_j\}}{1 + \exp \{\delta_j\} + C}$ where C is the sum of $\exp \{\delta_k\}$ for all products k different to j . Then, using the formula of the derivative of a quotient and taking into account that the derivative of function $\exp \{\delta_j\}$ with respect to δ_j is equal to $\exp \{\delta_j\}$ itself, we have that:

$$\begin{aligned} \frac{\partial s_j}{\partial \delta_j} &= \frac{\exp \{\delta_j\} [1 + \exp \{\delta_j\} + C]}{[1 + \exp \{\delta_j\} + C]^2} - \frac{\exp \{\delta_j\} \exp \{\delta_j\}}{[1 + \exp \{\delta_j\} + C]^2} \\ &= \frac{\exp \{\delta_j\}}{1 + \exp \{\delta_j\} + C} - \left[\frac{\exp \{\delta_j\}}{1 + \exp \{\delta_j\} + C} \right]^2 \\ &= s_j - (s_j)^2 = s_j(1 - s_j) \end{aligned}$$

Finally, combining $\frac{\partial s_j}{\partial \delta_j} = s_j(1 - s_j)$ and $\frac{\partial \delta_j}{\partial p_j} = -\alpha$, we have that:

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

(b) By the chain rule, $\frac{\partial s_j}{\partial p_k} = \frac{\partial s_j}{\partial \delta_k} \frac{\partial \delta_k}{\partial p_k}$ where $\delta_k = \beta x_k - \alpha p_k$. Note that $\frac{\partial \delta_k}{\partial p_k} = -\alpha$ such that $\frac{\partial s_j}{\partial p_k} = -\alpha \frac{\partial s_j}{\partial \delta_k}$. We need to obtain $\frac{\partial s_j}{\partial \delta_k}$. To obtain this (partial) derivative, we look at s_j as a function only of δ_k and we can treat all the δ' s for different to δ_k as constants. Therefore, we can write $s_j = \frac{A}{1 + \exp\{\delta_k\} + B}$ where A is $\exp\{\delta_j\}$ and B is the sum of $\exp\{\delta\}$ for all products different to k . Then, using the formula of the derivative of a quotient and taking into account that the derivative of function $\exp\{\delta_k\}$ with respect to δ_k is equal to $\exp\{\delta_k\}$ itself, we have that:

$$\begin{aligned} \frac{\partial s_j}{\partial \delta_k} &= -\frac{\exp\{\delta_k\} A}{[1 + \exp\{\delta_k\} + A]^2} \\ &= -\frac{\exp\{\delta_k\}}{1 + \exp\{\delta_k\} + A} \frac{A}{1 + \exp\{\delta_k\} + A} \\ &= -s_k s_j \end{aligned}$$

Finally, combining $\frac{\partial s_j}{\partial \delta_k} = -s_k s_j$ and $\frac{\partial \delta_k}{\partial p_k} = -\alpha$, we have that:

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

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) Show that the marginal conditions of profit maximization of firm A in this Bertrand game, $\frac{d\pi_A}{dp_{A1}} = 0$ and $\frac{d\pi_A}{dp_{A2}} = 0$, have the following form:

$$\begin{aligned} \frac{d\pi_A}{dp_{A1}} = 0 &\text{ implies } \frac{1}{\alpha} - (p_{A1} - c_{A1})(1 - s_{A1}) + (p_{A2} - c_{A2}) s_{A2} = 0 \\ \frac{d\pi_A}{dp_{A2}} = 0 &\text{ implies } \frac{1}{\alpha} + (p_{A1} - c_{A1}) s_{A1} - (p_{A2} - c_{A2})(1 - s_{A2}) = 0 \end{aligned}$$

ANSWER. The f.o.c. of profit maximization with respect to p_{A1} is:

$$q_{A1} + p_{A1} \frac{\partial q_{A1}}{\partial p_{A1}} - c_{A1} \frac{\partial q_{A1}}{\partial p_{A1}} + p_{A2} \frac{\partial q_{A2}}{\partial p_{A1}} - c_{A2} \frac{\partial q_{A2}}{\partial p_{A1}} = 0$$

or equivalently,

$$q_{A1} + (p_{A1} - c_{A1}) \frac{\partial q_{A1}}{\partial p_{A1}} + (p_{A2} - c_{A2}) \frac{\partial q_{A2}}{\partial p_{A1}} = 0$$

Since $q_j = H s_j$ and H is a constant, 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:

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

Dividing the two sides of this equation by αq_{A1} , we get:

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

Similarly, the f.o.c. of profit maximization with respect to p_{A2} is:

$$q_{A2} + p_{A2} \frac{\partial q_{A2}}{\partial p_{A2}} - c_{A2} \frac{\partial q_{A2}}{\partial p_{A2}} + p_{A1} \frac{\partial q_{A1}}{\partial p_{A2}} - c_{A1} \frac{\partial q_{A1}}{\partial p_{A2}} = 0$$

or equivalently,

$$q_{A2} + (p_{A2} - c_{A2}) \frac{\partial q_{A2}}{\partial p_{A2}} + (p_{A1} - c_{A1}) \frac{\partial q_{A1}}{\partial p_{A2}} = 0$$

Since $q_j = H s_j$ and H is a constant, we have that $\frac{\partial q_{A2}}{\partial p_{A2}} = H \frac{\partial s_{A2}}{\partial p_{A2}}$, and $\frac{\partial q_{A1}}{\partial p_{A2}} = H \frac{\partial s_{A1}}{\partial p_{A2}}$. 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 $\frac{\partial q_{A2}}{\partial p_{A2}} = -H \alpha s_{A2}(1 - s_{A2}) = -\alpha q_{A2}(1 - s_{A2})$, and $\frac{\partial q_{A1}}{\partial p_{A2}} = H \alpha s_{A1} s_{A2} = \alpha q_{A2} s_{A1}$. Therefore, this F.O.C. is:

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

Dividing the two sides of this equation by αq_{A2} , we get:

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

(b) Define the Price-Cost-Margins of products A1 and A2 as $PCM_{A1} \equiv p_{A1} - c_{A1}$ and $PCM_{A2} \equiv p_{A2} - c_{A2}$, respectively. In the First-Order-Conditions of profit

maximization in Q1.3(a), replace $p_{A1} - c_{A1}$ with PCM_{A1} , and $p_{A2} - c_{A2}$ with PCM_{A2} . Then, taking s_{A1} and s_{A2} as given, these First-Order-Conditions can be seen as a system of linear equation where the unknowns are PCM_{A1} and PCM_{A2} . Solve this system to obtain the following solution for PCM_{A1} and PCM_{A2} :

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

ANSWER. The system of equations is:

$$\begin{aligned} \frac{1}{\alpha} - PCM_{A1} (1 - s_{A1}) + PCM_{A2} s_{A2} &= 0 \\ \frac{1}{\alpha} + PCM_{A1} s_{A1} - PCM_{A2} (1 - s_{A2}) &= 0 \end{aligned}$$

We can use standard methods to solve this system of linear equations in the unknowns PCM_{A1} and PCM_{A2} . The following is a very simple one. If we subtract the second equation to the first, we get:

$$-PCM_{A1} + PCM_{A2} = 0$$

That is, for the profit maximizing pricing, the two price cost margins should be the same: $PCM_{A1} = PCM_{A2}$. Solving this restriction in any of the two equations, we get:

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

And solving for PCM_{A1} :

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

Q1.4. (20 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:

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

Or dividing both sides of the equation by αq_j , we get:

$$\frac{1}{\alpha} - (p_j - c_j) (1 - s_j) = 0$$

(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. (20 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. 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.

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 profits of the 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.

Note (for extra marks). The arguments above 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.

QUESTION 2. [120 points]. To answer the questions in this part of the problem set you need to use the dataset `cars1.dta`. Use this dataset to implement the estimations describe below.¹ 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 `Population/4`, i.e., `pop/4`
- Use prices measured in euros (`price`).
- For the product characteristics in the demand system, include the characteristics: horsepower, fuel, width, height, weight, domestic.
- Include also as explanatory variables the market characteristics: `log(pop)` and `log(gdp)`.
- In all the OLS estimations include fixed effects for market (country), year (year), and brand (brand).
- Include the price in levels (not in logarithms).

Q2.1. (70 points)

(a) [10 points] Obtain the OLS-Fixed effects estimator of the Standard logit model. Interpret the results.

ANSWER. First, we read the dataset and construct some new variables.

```
use "C:\cars1.dta", clear
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
```

¹This is the same dataset that Frank Verboven provides for his Stata command `mergesim`. Note that this is a different sample than the one we have used in the Tutorials.


```

egen sum_share = sum(share), by(country year)
gen share0 = 1 - sum_share
//generate log odd ratio
gen lsj_ls0 = ln(share/share0)

```

And we obtain the Fixed Effects estimator of the Standard Logit. To implement this Fixed effects estimator, we could use command `"reghdfe , a(country year brand)"`. That is perfectly correct. Instead, here I have used the standard OLS command `"reg"` and I have included explicitly the dummy variables for country year and brand. Of course, that is exactly (by definition) the fixed effects with country, year, and brand fixed effects. We use the `reghdfe` command when the number of fixed effects is very large (hundreds, thousands, or more) such that from a computational point of view it is much more time efficient to implement the FE estimator using the within-groups transformation of the model. When the number of fixed effects is relatively small (less than a few dozens) we can implement the FE estimator by using the standard OLS command `"reg"` including the dummy variables explicitly. Note that we do not need to generate the dummy variables before using the command `"reg"`. We can easily include them by using `i.country i.year i.brand`.

```

//      Fixed Effects estimation Standard Logit
reg lsj_ls0 price horsepower fuel width domestic height weight logpop loggdp
i.country i.year i.brand, vce(robust)

```

This is the table with the estimation results. I have "cut" the table to avoid the whole list of coefficient for the year and brand fixed effects.

Interpretation of the estimation results.

- The parameter for **price** (i.e., minus the marginal utility of income) is significantly smaller than zero, which is consistent with economic theory.

- The estimates for the parameters associated to product characteristics (i.e., the marginal utilities of these product attributes) are all significantly different than zero. Some of these estimates have plausible signs but others do not. The negative marginal utility of the characteristic **fuel** consumption is very plausible, and so are the positive marginal utilities of the characteristics **domestic** and **width**. There are also arguments to find plausible the negative estimates for the marginal utilities of **height** and **weight**, e.g., taking **width** as given, an increase in **height** and **weight** can reduce performance, safety, and services

```
. reg lsj_ls0 price horsepower fuel width domestic height weight logpop loggdp i.country i.year i.brand, vce
> (robust)
```

```
Linear regression                               Number of obs =   11483
                                                F(   79, 11403) =   217.90
                                                Prob > F       =    0.0000
                                                R-squared      =    0.5739
                                                Root MSE     =    .98332
```

lsj_ls0	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
price	-.0540465	.0044622	-12.11	0.000	-.0627931	-.0452998
horsepower	-.016989	.0015023	-11.31	0.000	-.0199338	-.0140441
fuel	-.0719452	.0131289	-5.48	0.000	-.09768	-.0462103
width	.0533045	.0024859	21.44	0.000	.0484317	.0581772
domestic	1.78407	.0265343	67.24	0.000	1.732058	1.836082
height	-.0146662	.002537	-5.78	0.000	-.0196392	-.0096932
weight	-.0008459	.0001628	-5.20	0.000	-.001165	-.0005268
logpop	.7517001	.1955416	3.84	0.000	.3684048	1.134995
loggdp	.1070581	.0503965	2.12	0.034	.0082722	.2058439
country						
2	-2.164682	.3350884	-6.46	0.000	-2.821513	-1.507851
3	-2.135312	.3747761	-5.70	0.000	-2.869937	-1.400686
4	-2.540298	.4514167	-5.63	0.000	-3.425153	-1.655444
5	-1.453818	.3390208	-4.29	0.000	-2.118357	-.7892787

of a car. However, the estimated negative marginal utility of **horsepower** does not seem plausible.

- There are several interpretations for this implausible estimate of the coefficient for **horsepower**. Let me include here two possible explanations. First, consumers may have heterogeneous taste for this characteristic. Some consumers like it and others dislike them. We are estimating just the average taste. Though this is possible, it is still hard to believe that the average consumer valuation of **horsepower** is negative. A second possible explanation is that **horsepower** is negatively correlated with the unobserved quality of the product, ξ_j . For instance, it could be negatively correlated with the unobserved safety of a car. The OLS estimate of the marginal utility of **horsepower** is capturing both the direct (ceteris paribus) effect of this attribute on consumer utility and also the indirect (non causal) effect because its negative correlation with car safety that is part of the error term.

- The effect of **loggdp** on the demand of cars is positive and statistically significant. Markets and time periods with more **loggdp** have a larger demand for all the products. This positive income effect makes economic sense.

- The effect of **logpop** on the demand of cars is also positive and statistically significant. Markets/years with more population tend to have consumers with a higher taste for cars. This effect can be also interpreted as a correction for our measure of market size, $\text{pop}/4$.

(b) [20 points] Obtain the Instrumental Variables (IV) estimator of the Standard logit model. Use the following set of instruments. For each product characteristic (horsepower, fuel, width, height, weight, and domestic.) we can construct three instruments described below.

Let j index car model (variable **co**), let m index country (variable **country**), and

let t index year (variable year). In this dataset we have 6 product characteristics: horsepower, fuel, width, height, weight, and domestic. For the sake of concreteness, I focus here on the construction of the instrumental variables associated with one of these product characteristics, say fuel. Note that fuel_{jmt} represents observation (j, m, t) of variable fuel. Let IV1_fuel , IV2_fuel , and IV3_fuel be the three instrumental variables associated with fuel. IV1_fuel_{jmt} represents observation (j, m, t) of variable IV1_fuel . Similarly, IV2_fuel_{jmt} and IV3_fuel_{jmt} represent observation (j, m, t) of variables IV2_fuel and IV3_fuel , respectively. These instruments are defined as follows:

$$\begin{aligned}\text{IV1_fuel}_{jmt} &= \sum_{i=1}^{J_{mt}} (\text{fuel}_{jmt} - \text{fuel}_{imt}) \\ \text{IV2_fuel}_{jmt} &= \sum_{i=1}^{J_{mt}} |\text{fuel}_{jmt} - \text{fuel}_{imt}| \\ \text{IV3_fuel}_{jmt} &= \sum_{i=1}^{J_{mt}} (\text{fuel}_{jmt} - \text{fuel}_{imt})^2\end{aligned}$$

where J_{mt} is the number of car models in the dataset in country m at year t .

For your IV estimator, use the 12 instruments from IV1 and IV2 or from IV1 and IV3. Interpret the results from this IV estimation.

ANSWER. We first construct the instruments. Here I present the construction of instruments IV1 and IV3. For the construction of IV2, see the Francis Guiton's notes on the tutorial of how to use loops in Stata to construct variables such as instrument IV2.

[i] STATA Code to construct IV1s. Note that:

$$\begin{aligned}\text{IV1_fuel}_{jmt} &= \sum_{i=1}^{J_{mt}} (\text{fuel}_{jmt} - \text{fuel}_{imt}) \\ &= J_{mt} * \text{fuel}_{jmt} - \sum_{i=1}^{J_{mt}} \text{fuel}_{imt} \\ &= J_{mt} * \text{fuel}_{jmt} - \text{sumfuel}_{mt}\end{aligned}$$

where sumfuel_{mt} is $\sum_{i=1}^{J_{mt}} \text{fuel}_{imt}$. Based on this formula, we can construct the IV1 variables using the following lines of code in STATA.

```

// constructing instruments IV1
egen numJ = sum(1), by(country year)
egen sumsome = sum(horsepower), by(country year)
gen IV1_horsepower = numJ * horsepower - sumsome
drop sumsome

egen sumsome = sum(fuel), by(country year)
gen IV1_fuel = numJ * fuel - sumsome
drop sumsome

egen sumsome = sum(width), by(country year)
gen IV1_width = numJ * width - sumsome
drop sumsome

egen sumsome = sum(domestic), by(country year)
gen IV1_domestic = numJ * domestic - sumsome
drop sumsome

egen sumsome = sum(height), by(country year)
gen IV1_height = numJ * height - sumsome
drop sumsome

egen sumsome = sum(weight), by(country year)
gen IV1_weight = numJ * weight - sumsome
drop sumsome

```

[ii] STATA Code to construct IV3s. Define meanfuel_{mt} as the mean value $\frac{1}{J_{mt}} \sum_{i=1}^{J_{mt}} \text{fuel}_{imt}$, where J_{mt} is the number of car models in the dataset in country m at year t . It is simple to verify that $\sum_{j=1}^{J_{mt}} (\text{fuel}_{jmt} - \text{meanfuel}_{mt}) = 0$, and we will use this result below. By adding and subtracting meanfuel_{mt} in the expression for $IV3_fuel_{jmt}$, we can get the following

result:

$$\begin{aligned}
IV3_fuel_{jmt} &= \sum_{i=1}^{J_{mt}} ([fuel_{jmt} - meanfuel_{mt}] - [fuel_{imt} - meanfuel_{mt}])^2 \\
&= \sum_{i=1}^{J_{mt}} [fuel_{jmt} - meanfuel_{mt}]^2 \\
&\quad - 2 \sum_{i=1}^{J_{mt}} [fuel_{jmt} - meanfuel_{mt}] [fuel_{imt} - meanfuel_{mt}] \\
&\quad + \sum_{i=1}^{J_{mt}} [fuel_{imt} - meanfuel_{mt}]^2 \\
&= J_{mt} * [fuel_{jmt} - meanfuel_{mt}]^2 \\
&\quad - 2[fuel_{jmt} - meanfuel_{mt}] \sum_{i=1}^{J_{mt}} [fuel_{imt} - meanfuel_{mt}] \\
&\quad + J_{mt} * varfuel_{mt} \\
&= J_{mt} * [fuel_{jmt} - meanfuel_{mt}]^2 + J_{mt} * varfuel_{mt}
\end{aligned}$$

where $varfuel_{mt}$ is the variance $\frac{1}{J_{mt}} \sum_{i=1}^{J_{mt}} [fuel_{imt} - meanfuel_{mt}]^2$, and we have used the result $\sum_{j=1}^{J_{mt}} (fuel_{jmt} - meanfuel_{mt}) = 0$.

Using this formula, we can construct the IV3 variables using the following lines of code in STATA. Note that the **egen** command does not admit the `variance(.)` function but it admits the standard deviation function, `sd(.)`.

```
// constructing instruments IV3
egen meansome = mean(horsepower), by(country year)
egen sdsome = sd(horsepower), by(country year)
gen IV3_horsepower = numJ * (horsepower - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

egen meansome = mean(fuel), by(country year)
egen sdsome = sd(fuel), by(country year)
gen IV3_fuel = numJ * (fuel - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

egen meansome = mean(width), by(country year)
egen sdsome = sd(width), by(country year)
gen IV3_width = numJ * (width - meansome)^2 + numJ * (sdsome^2)
```

```

drop meansome sdsome

egen meansome = mean(domestic), by(country year)
egen sdsome = sd(domestic), by(country year)
gen IV3_domestic = numJ * (domestic - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

egen meansome = mean(height), by(country year)
egen sdsome = sd(height), by(country year)
gen IV3_height = numJ * (height - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

egen meansome = mean(weight), by(country year)
egen sdsome = sd(weight), by(country year)
gen IV3_weight = numJ * (weight - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

```

[iii] STATA code to implement the IV estimator. To implement this IV estimator, we could use command `"ivreghdfe , a(country year brand)"`. That is perfectly correct. Instead, here I have used the standard IV command in Stata, `"ivreg2"` and I have included explicitly the dummy variables for `country`, `year`, and `brand`. My argument for doing this is exactly the same as above for implementing the FE estimator using `"reg"` instead of `"reghdfe"`, i.e., the number of fixed effects (dummy variables) in this estimation is quite small.

This is the Stata code for this IV estimation.

```

// IV estimator of Standard Logit model with fixed effects
ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country
i.year i.brand (price = IV*), robust

```

This is the table with the estimation results. I have "cut" the table to avoid the whole list of coefficients for the year and brand fixed effects. I will include later the part of the table for the over-identifying restrictions.

Interpretation of the IV estimation results.

- The parameter for `price` (i.e., minus the marginal utility of income) is significantly smaller than zero, which is consistent with economic theory. Interestingly, the IV estimate of this parameter is not smaller than the FE estimate presented above. This is not consistent

```
. ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country i.year i.brand (price
> = IV*), robust
```

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

		Number of obs =	11483
		F(79, 11403) =	217.63
		Prob > F =	0.0000
		Centered R2 =	0.5733
		Uncentered R2 =	0.9824
		Root MSE =	.9806
Total (centered) SS	=	25875.093	
Total (uncentered) SS	=	627916.7574	
Residual SS	=	11041.62661	

lsj_ls0	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
price	-.0420901	.0069636	-6.04	0.000	-.0557386	-.0284416
horsepower	-.0200002	.0021292	-9.39	0.000	-.0241734	-.015827
fuel	-.0728848	.0130869	-5.57	0.000	-.0985346	-.0472349
width	.0540433	.0024751	21.83	0.000	.0491921	.0588945
domestic	1.785494	.0264425	67.52	0.000	1.733667	1.83732
height	-.0149621	.0025384	-5.89	0.000	-.0199372	-.009987
weight	-.0009512	.0001635	-5.82	0.000	-.0012716	-.0006308
logpop	.6170441	.2060711	2.99	0.003	.2131522	1.020936
loggdp	.1328915	.0515467	2.58	0.010	.0318619	.2339212

with the typical interpretation that **price** is positively correlated with the unobserved quality of the product and this generates an upward bias in the OLS/FE estimator. It seems that the "correction" introduced by this IV estimator is not going in that direction. We can interpret this result in two possible ways. The "optimistic" interpretation is that the instruments we are using are valid, but the OLS/FE estimator was already unbiased/consistent such that there are not significant differences in the parameter estimates using OLS/FE and using IV. The "pessimistic" interpretation is that at least one of the instruments we are using is not valid – it is correlated with the error term – such that the IV estimator is not correction for the bias in the OLS/FE estimator. The test of over-identifying restrictions will help us to choose between these two interpretations, and in particular, it will provide evidence in support of the "pessimistic" one. See below.

- For the rest of the parameter estimates, their signs are exactly the same as in the OLS/FE estimation, and their magnitude are also very similar. Therefore, the same comments as for the OLS/FE estimation apply here.

- The standard errors of the estimated coefficient for the **price** variable is almost twice in the IV estimation than in the OLS/FE. This is always the case. The IV estimation does not use all the sample variation of the endogenous **price** variable (only the sample variation explained by the instruments) and this always implies a loss of precision in the parameter estimate.

(c) [10 points] Using the IV estimation, obtain how much the average consumer is willing to pay (in Euros) for a reduction of 1 unit in the characteristic fuel (that is, for an improvement in fuel efficiency of 1 liter per km).

ANSWER. A consumer (maximum) willingness to pay for a unit change in a product characteristic is equal to the change in the price of the product that leaves the consumer indifferent between the following two scenarios: "no change in the product characteristic and no change in price" and "one unit change in the product characteristic and change in price".

Given a (indirect) utility function $-\alpha p_j + \beta_1 x_{1j} + \dots + \beta_K x_{Kj} + \xi_j$, let's obtain the expression for the willingness to pay (WTP) for a unit change in product characteristic x_1 . Using the definition in the previous paragraph, we have that the willingness to pay is the change in price Δ that solves the following equation:

$$-\alpha p_j + \beta_1 x_{1j} + \dots + \beta_K x_{Kj} + \xi_j = -\alpha[p_j + \Delta] + \beta_1[x_{1j} + 1] + \dots + \beta_K x_{Kj} + \xi_j$$

Solving for Δ in this equation, we have that:

$$WTP = \Delta = \frac{\beta_1}{\alpha}$$

Since β_1 is measured in "utils/unit of x_1 " and α is measured in "utils/dollars" (if price is measured in dollars), then we have that WTP is measured in dollars/unit of x_1 .

Now, let us consider the specific question. Variable **fuel** is measured in liters if fuel consumption per km. The IV estimate of the parameter β_{fuel} is -0.072 utils/liter per km. Variable **price** is measured in thousands of euros. The IV estimate of the parameter α is 0.042 utils/1000 euros. Therefore, the WTP for a reduction in one unit in the fuel variable is:

$$WTP = (-1) \frac{\beta_{\text{fuel}}}{\alpha} = \frac{0.072 \text{ utils/ liter per km}}{0.042 \text{ utils/1000 euros}} = 1,714 \text{ euros (of year 1999)}$$

Is this estimate of WTP reasonable? We can calculate the savings that this improvement in fuel efficiency implies for an average driver. The average driving mileage of a car in UK is 7,400 miles (11,900 km) a year. The average price of gasoline in UK in year 1999 was 0.866 euros per liter (0.619 £ per liter \times 1.40 euros/£). Therefore, an improvement in fuel efficiency of 1 liter per km implies annual savings of $11,900 \text{ liter} \times 0.866 \text{ euros} = 10,305$ euros. Therefore, it seems that the estimated WTP is too small relative to the savings it implies for the average driver.

(d) [10 points] Test the over-identifying restrictions in the IV estimation.

ANSWER. Stata `ivreg2` command reports Hansen-Sargan test of over-identifying restrictions. The null hypothesis is that the instruments are not correlated with the error term. This is the result of this test:


```
-----  
Hansen J statistic (overidentification test of all instruments): 342.970  
Chi-sq(11) P-val = 0.0000  
-----
```

The p-value is very close to zero. There is strong evidence against the validity of the instruments.

(e) [10 points] Using the IV estimation, test the null hypothesis that all countries have the same price coefficient.

ANSWER. We need to implement a new IV estimator that allows the price coefficient to be different across countries. We need to interact the `price` variable and the instrumental variables with country dummies. We first generate the variables with these interactions.

```
// Generating interactions of price and country dummies  
gen price_dumBel = price * (country==1)  
gen price_dumFra = price * (country==2)  
gen price_dumGer = price * (country==3)  
gen price_dumIta = price * (country==4)  
gen price_dumUK = price * (country==5)  
  
// Generating interactions of IV1 and country dummies  
gen IV1_dumBel_horsepower = IV1_horsepower * (country==1)  
gen IV1_dumBel_fuel = IV1_fuel * (country==1)  
gen IV1_dumBel_width = IV1_width * (country==1)  
gen IV1_dumBel_domestic = IV1_domestic * (country==1)  
gen IV1_dumBel_height = IV1_height * (country==1)  
gen IV1_dumBel_weight = IV1_weight * (country==1)  
  
gen IV1_dumFra_horsepower = IV1_horsepower * (country==2)  
gen IV1_dumFra_fuel = IV1_fuel * (country==2)  
gen IV1_dumFra_width = IV1_width * (country==2)  
gen IV1_dumFra_domestic = IV1_domestic * (country==2)  
gen IV1_dumFra_height = IV1_height * (country==2)  
gen IV1_dumFra_weight = IV1_weight * (country==2)  
  
gen IV1_dumGer_horsepower = IV1_horsepower * (country==3)  
gen IV1_dumGer_fuel = IV1_fuel * (country==3)
```

```

gen IV1_dumGer_width = IV1_width * (country==3)
gen IV1_dumGer_domestic = IV1_domestic * (country==3)
gen IV1_dumGer_height = IV1_height * (country==3)
gen IV1_dumGer_weight = IV1_weight * (country==3)

gen IV1_dumIta_horsepower = IV1_horsepower * (country==4)
gen IV1_dumIta_fuel = IV1_fuel * (country==4)
gen IV1_dumIta_width = IV1_width * (country==4)
gen IV1_dumIta_domestic = IV1_domestic * (country==4)
gen IV1_dumIta_height = IV1_height * (country==4)
gen IV1_dumIta_weight = IV1_weight * (country==4)

gen IV1_dumUK_horsepower = IV1_horsepower * (country==5)
gen IV1_dumUK_fuel = IV1_fuel * (country==5)
gen IV1_dumUK_width = IV1_width * (country==5)
gen IV1_dumUK_domestic = IV1_domestic * (country==5)
gen IV1_dumUK_height = IV1_height * (country==5)
gen IV1_dumUK_weight = IV1_weight * (country==5)

// Generating interactions of IV3 and country dummies
gen IV3_dumBel_horsepower = IV3_horsepower * (country==1)
gen IV3_dumBel_fuel = IV3_fuel * (country==1)
gen IV3_dumBel_width = IV3_width * (country==1)
gen IV3_dumBel_domestic = IV3_domestic * (country==1)
gen IV3_dumBel_height = IV3_height * (country==1)
gen IV3_dumBel_weight = IV3_weight * (country==1)

gen IV3_dumFra_horsepower = IV3_horsepower * (country==2)
gen IV3_dumFra_fuel = IV3_fuel * (country==2)
gen IV3_dumFra_width = IV3_width * (country==2)
gen IV3_dumFra_domestic = IV3_domestic * (country==2)
gen IV3_dumFra_height = IV3_height * (country==2)
gen IV3_dumFra_weight = IV3_weight * (country==2)

gen IV3_dumGer_horsepower = IV3_horsepower * (country==3)
gen IV3_dumGer_fuel = IV3_fuel * (country==3)
gen IV3_dumGer_width = IV3_width * (country==3)
gen IV3_dumGer_domestic = IV3_domestic * (country==3)
gen IV3_dumGer_height = IV3_height * (country==3)
gen IV3_dumGer_weight = IV3_weight * (country==3)

```

```

. // IV estimator of model with heterogeneous price parameters
. ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country i.year i.brand (price_
> dum* = IV1_dum* IV3_dum*), robust
Warning - collinearities detected
Vars dropped:          IV1_dumBel_domestic IV3_dumBel_domestic

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

Total (centered) SS      =    25875.093
Total (uncentered) SS   =   627916.7574
Residual SS             =   11012.41094

Number of obs =    11483
F( 83, 11399) =   208.98
Prob > F       =    0.0000
Centered R2    =    0.5744
Uncentered R2  =    0.9825
Root MSE      =    .9793

```

lsj_ls0	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
price_dumBel	-.0444721	.0066609	-6.68	0.000	-.0575272	-.031417
price_dumFra	-.0494017	.0071555	-6.90	0.000	-.0634262	-.0353772
price_dumGer	-.0465184	.0084123	-5.53	0.000	-.0630062	-.0300306
price_dumIta	-.0524841	.0067833	-7.74	0.000	-.0657791	-.0391892
price_dumUK	-.0300085	.0054395	-5.52	0.000	-.0406696	-.0193473
horsepower	-.0196524	.0020211	-9.72	0.000	-.0236136	-.0156912
fuel	-.0728265	.0129441	-5.63	0.000	-.0981965	-.0474565
width	.0544559	.0024918	21.85	0.000	.0495721	.0593397
domestic	1.787201	.0270579	66.05	0.000	1.734168	1.840233
height	-.015331	.0025132	-6.10	0.000	-.0202567	-.0104052
weight	-.000954	.000162	-5.89	0.000	-.0012715	-.0006365
logpop	.6391779	.2062862	3.10	0.002	.2348643	1.043492
loggdp	.111926	.0514538	2.18	0.030	.0110784	.2127735

```

gen IV3_dumIta_horsepower = IV3_horsepower * (country==4)
gen IV3_dumIta_fuel = IV3_fuel * (country==4)
gen IV3_dumIta_width = IV3_width * (country==4)
gen IV3_dumIta_domestic = IV3_domestic * (country==4)
gen IV3_dumIta_height = IV3_height * (country==4)
gen IV3_dumIta_weight = IV3_weight * (country==4)

gen IV3_dumUK_horsepower = IV3_horsepower * (country==5)
gen IV3_dumUK_fuel = IV3_fuel * (country==5)
gen IV3_dumUK_width = IV3_width * (country==5)
gen IV3_dumUK_domestic = IV3_domestic * (country==5)
gen IV3_dumUK_height = IV3_height * (country==5)
gen IV3_dumUK_weight = IV3_weight * (country==5)

```

Then, we implement the IV estimator.

```

// IV estimator of model with heterogeneous price parameters
ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country
i.year i.brand (price_dum* = IV1_dum* IV3_dum*), robust

```

This is the table of estimation results:

The estimated parameters are quite similar to the ones from the model that restricts the price coefficient to be the same across countries. However, there is some heterogeneity across countries. In particular, consumers in UK are less price sensitive than in the other countries.

```

. // Testing the null hypothesis of homogeneous price coefficients
. test price_dumBel = price_dumFra = price_dumGer = price_dumIta = price_dumUK

( 1)  price_dumBel - price_dumFra = 0
( 2)  price_dumBel - price_dumGer = 0
( 3)  price_dumBel - price_dumIta = 0
( 4)  price_dumBel - price_dumUK = 0

      chi2( 4) =    38.79
    Prob > chi2 =    0.0000

. // Similar test but excluding UK
. test price_dumBel = price_dumFra = price_dumGer = price_dumIta

( 1)  price_dumBel - price_dumFra = 0
( 2)  price_dumBel - price_dumGer = 0
( 3)  price_dumBel - price_dumIta = 0

      chi2( 3) =     4.31
    Prob > chi2 =    0.2300

```

Finally, we test the null hypothesis that the price coefficients are the same across countries. This is the code line.

```

// Testing the null hypothesis of homogeneous price coefficients
test price_dumBel = price_dumFra = price_dumGer = price_dumIta = price_dumUK

```

And this is the table with the result of this test.

The p-value of the test is very close to zero, so there is strong evidence against the homogeneity of the price coefficients. Since this rejection seems to be associated to UK, it is relevant to test a similar null hypothesis but excluding UK. This is the result:

The p-value of 0.23 such that we cannot reject this null hypothesis under the usual significance levels.

(f) [10 points] Based on the IV estimation, obtain the average price elasticity of demand evaluated at the mean values of prices and market shares.

ANSWER. The elasticity of demand in observation (j, m, t) is: $\frac{\partial s_{jmt}}{\partial p_{jmt}} \frac{p_{jmt}}{s_{jmt}} = -\alpha_m s_{jmt} (1 - s_{jmt}) \frac{p_{jmt}}{s_{jmt}}$, where α_m is the price coefficient for country m .

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

In Stata, after an estimation command, the estimated coefficient associated with a variable with name `varname` is saved in variable called `_b[varname]`. Therefore, the price coefficients are stored in `_b[price_dumBel]`, `_b[price_dumFra]`, `_b[price_dumGer]`, `_b[price_dumIta]`,

```
. sum elasticity , detail
```

elasticity				
	Percentiles	Smallest		
1%	-2.045456	-5.363294		
5%	-1.500102	-4.726171		
10%	-1.29197	-4.5112	Obs	11483
25%	-.9690685	-4.495609	Sum of Wgt.	11483
50%	-.708903		Mean	-.8006961
		Largest	Std. Dev.	.3777577
75%	-.5367818	-.2281124		
90%	-.4376537	-.2271975	Variance	.1427009
95%	-.3927333	-.2256188	Skewness	-2.082282
99%	-.3156107	-.2242705	Kurtosis	12.6209

and `_b[price_dumUK]`. This is the code line to generate the variable with the elasticity at each observation.

```
// Calculating elasticities
gen elasticity = (1-share)*(price_dumBel*_b[price_dumBel] + price_dumFra*_b[price_dumFra]
+ price_dumGer*_b[price_dumGer] + price_dumIta*_b[price_dumIta] + price_dumUK*_b[price_dumUK])
```

To obtain the average elasticity, we just obtain the mean of variable `elasticity` using the command `summarize`. To see also the median and percentiles in the distribution, we use `summarize, detail`.

The mean value of the elasticity is -0.70 and the median is -0.80 . More than 75% of the observations have elasticities smaller greater than -1 . This is values for the elasticity are too low. This result seems to indicate that we have not been able to control well enough for endogeneity such that our estimate of the price coefficient is upward biased (biased towards zero).

Q2.2. (50 points) Consider the equilibrium condition (first order conditions of profit maximization) under the assumption that each product is produced by only one firm.

(a) (5 points) Write the equation for this equilibrium condition. Using this condition, obtain an estimate of the (realized) marginal cost for every car-market-year observation in the data.

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 MC_j , we have that:

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

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

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

Plugging this expression into the previous equation for MC_j , we have:

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

This is the Stata code to obtain the realized marginal cost at every observation (j, m, t) .

```
//    Calculating realized marginal cost
gen mcost = .
replace mcost = price + 1/(_b[price_dumBel]*(1-share)) if country==1
replace mcost = price + 1/(_b[price_dumFra]*(1-share)) if country==2
replace mcost = price + 1/(_b[price_dumGer]*(1-share)) if country==3
replace mcost = price + 1/(_b[price_dumIta]*(1-share)) if country==4
replace mcost = price + 1/(_b[price_dumUK] *(1-share)) if country==5
```

Taking a look at the distribution of `mcost` (using `summarize, detail`) we can see that for more than 75% of the observations the marginal cost is negative. Again, this is a direct implication of the very small estimate of the price coefficients that measure the price sensitivity of demand. Due to endogeneity problems, the estimate of α is biased towards zero, this implies that the term $\frac{1}{\alpha(1-s_j)}$ is too large, and it generates negative values for the estimate of MC_j .

(b) (15 points) 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 horsepower, fuel, width, height, weight, domestic. Interpret the estimated coefficients.

```
. reg mcost horsepower fuel width domestic height weight i.country i.year i.brand, vce(robust)
```

Linear regression

Number of obs = 11483
F(76, 11406) = 649.23
Prob > F = 0.0000
R-squared = 0.8682
Root MSE = 3.2238

mcost	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
horsepower	.2540999	.0077438	32.81	0.000	.2389208	.269279
fuel	.0687048	.0542524	1.27	0.205	-.0376392	.1750488
width	-.0633682	.0073867	-8.58	0.000	-.0778474	-.0488889
domestic	-.1961678	.0857187	-2.29	0.022	-.3641911	-.0281445
height	.0250448	.007183	3.49	0.000	.0109649	.0391247
weight	.0087744	.0007252	12.10	0.000	.0073529	.0101959

ANSWER. This is the Stata code

```
// Fixed Effects estimation of Marginal Cost function
reg mcost horsepower fuel width domestic height weight i.country i.year i.brand,
vce(robust)
```

And this is the table of estimation results.

Interpretation of estimated parameters

- The sign of the estimated coefficients for **horsepower** and **domestic** have economic sense and they are statistically significant. It is more costly to produce a car with more power, and it is cheaper to produce (and distribute) a car that is sold domestically. The coefficient for **domestic** implies that ceteris paribus is 1,961 euros cheaper to produce a car domestically.

- However, the estimated coefficient for **fuel** (inefficiency) is not plausible. It says that it is more costly to produce a car that is less fuel efficient, or equivalently, that improving the fuel efficiency of a car implies a reduction in its marginal cost. This does not have economic sense. However, this estimated coefficient is not significantly different to zero.

(c) (20 points) Now consider that the marginal cost may also depend on the amount produced and sold of the product (variable **qu**). Include this variable in the regression for the marginal cost function. Estimate this marginal cost function by Instrumental variables using the same instruments as for the demand estimation, i.e., the characteristics of competing products. Interpret the estimated coefficients.

IV (2SLS) estimation

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity

		Number of obs =	11483
		F(77, 11405) =	625.57
		Prob > F =	0.0000
Total (centered) SS	=	Centered R2	= 0.8605
Total (uncentered) SS	=	Uncentered R2	= 0.8937
Residual SS	=	Root MSE	= 3.306

mcost	Robust		z	P> z	[95% Conf. Interval]	
	Coef.	Std. Err.				
qu	.0000238	5.07e-06	4.70	0.000	.0000139	.0000338
horsepower	.261996	.0087405	29.98	0.000	.244865	.279127
fuel	.0925518	.056399	1.64	0.101	-.0179882	.2030918
width	-.0804578	.0088229	-9.12	0.000	-.0977504	-.0631652
domestic	-1.483163	.2967525	-5.00	0.000	-2.064787	-.9015387
height	.0314073	.0074228	4.23	0.000	.0168588	.0459558
weight	.0093799	.0006933	13.53	0.000	.008021	.0107387

ANSWER. This is the Stata code.

```
//      IV estimation of Marginal Cost function including qu
ivreg2 mcost horsepower fuel width domestic height weight i.country i.year
i.brand (qu = IV1_dum* IV3_dum*), robust
```

This is the table with the estimation results.

Interpretation of estimated parameters

- The estimated coefficients for the product characteristics are very similar as in the OLS/FE estimation. The same comments apply.

- The estimated coefficient for the amount of output is significantly greater than zero. There is evidence of increasing marginal cost with output.

(d) (10 points) Based on the IV estimation of the marginal cost function in Q2.2(c), obtain the change in the marginal cost in Euros of a change in output of when a firm increases in 10,000 units (cars per market and year) the variable qu. Is this estimate plausible? Discuss this result.

ANSWER. Variable **qu** is measure in number of cars. The marginal cost is measured in the same units as **price**, in thousands of euros. Since the estimated coefficient for variable **qu** is 0.0000238, this implies that 1 more car increases the marginal cost in 0.0238 euros, that

is approximately 2 cents of euro. Therefore, an increase in 10,000 cars in `qu` implies an increase in the marginal cost of 238 euros.

Is this plausible? This seems plausible because it is positive (a decreasing marginal cost function is very implausible) and because it is not a large increase. In fact, it is a very modest increase that indicates that the marginal cost is almost constant.

Why we can say that a 238 euros increase in marginal cost for 10,000 more cars is a modest increase? Looking at the empirical distribution of variable `qu` (using `summarize, detail`), we can see that a change in output in 10,000 cars is a very substantial increase. The median of `qu` is 6262 cars and the percentile 75% is 18,868. In contrast, 238 euros represents a small fraction of a car price. The median price is 16,400 euros such that 238 euros represents only 1.5% of this median price.

STATA CODE

```
// -----
//   EC0310 - Problem set #2
//
// Data on demand of differentiated products
// Verboven's data on automobiles in Europe
//
// Victor Aguirregabiria
//   March 15, 2021
//-----
// -----
//   Reading data and constructing new variables
// -----
log using "C:\problemset_02_2021.log", replace
use "C:\cars1.dta", clear
// -----
//   Question 2.1(a) Fixed Effects estimation Standard Logit
// -----
// constructing some variables
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(country year)
gen share0 = 1 - sum_share
sum share share0
//generate log odd ratio
gen lsj_ls0 = ln(share/share0)
//   Fixed Effects estimation Standard Logit
reg lsj_ls0 price horsepower fuel width domestic height weight logpop loggdp
i.country i.year i.brand, vce(robust)
// -----
//   Question 2.1(b) IV estimation Standard Logit
```

```

// -----
// constructing instruments IV1
egen numJ = sum(1), by(country year)
egen sumsome = sum(horsepower), by(country year)
gen IV1_horsepower = numJ * horsepower - sumsome
drop sumsome
egen sumsome = sum(fuel), by(country year)
gen IV1_fuel = numJ * fuel - sumsome
drop sumsome
egen sumsome = sum(width), by(country year)
gen IV1_width = numJ * width - sumsome
drop sumsome
egen sumsome = sum(domestic), by(country year)
gen IV1_domestic = numJ * domestic - sumsome
drop sumsome
egen sumsome = sum(height), by(country year)
gen IV1_height = numJ * height - sumsome
drop sumsome
egen sumsome = sum(weight), by(country year)
gen IV1_weight = numJ * weight - sumsome
drop sumsome
sum IV1*
corr IV*
// constructing instruments IV3
egen meansome = mean(horsepower), by(country year)
egen sdsome = sd(horsepower), by(country year)
gen IV3_horsepower = numJ * (horsepower - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome
egen meansome = mean(fuel), by(country year)
egen sdsome = sd(fuel), by(country year)
gen IV3_fuel = numJ * (fuel - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome
egen meansome = mean(width), by(country year)
egen sdsome = sd(width), by(country year)
gen IV3_width = numJ * (width - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome

```

```

egen meansome = mean(domestic), by(country year)
egen sdsome = sd(domestic), by(country year)
gen IV3_domestic = numJ * (domestic - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome
egen meansome = mean(height), by(country year)
egen sdsome = sd(height), by(country year)
gen IV3_height = numJ * (height - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome
egen meansome = mean(weight), by(country year)
egen sdsome = sd(weight), by(country year)
gen IV3_weight = numJ * (weight - meansome)^2 + numJ * (sdsome^2)
drop meansome sdsome
sum IV3*
corr IV3*
// IV estimator of Standard Logit model with fixed effects
ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country
i.year i.brand (price = IV*), robust
predict ivres, resid
reg ivres horsepower fuel width domestic height weight logpop loggdp i.country
i.year i.brand IV*
// -----
// Question 2.1(e) IV estimation with heterogeneous price coefficients
// -----
// Generating interactions of price and country dummies
gen price_dumBel = price * (country==1)
gen price_dumFra = price * (country==2)
gen price_dumGer = price * (country==3)
gen price_dumIta = price * (country==4)
gen price_dumUK = price * (country==5)
// Generating interactions of IV1 and country dummies
gen IV1_dumBel_horsepower = IV1_horsepower * (country==1)
gen IV1_dumBel_fuel = IV1_fuel * (country==1)
gen IV1_dumBel_width = IV1_width * (country==1)
gen IV1_dumBel_domestic = IV1_domestic * (country==1)
gen IV1_dumBel_height = IV1_height * (country==1)
gen IV1_dumBel_weight = IV1_weight * (country==1)

```

```

gen IV1_dumFra_horsepower = IV1_horsepower * (country==2)
gen IV1_dumFra_fuel = IV1_fuel * (country==2)
gen IV1_dumFra_width = IV1_width * (country==2)
gen IV1_dumFra_domestic = IV1_domestic * (country==2)
gen IV1_dumFra_height = IV1_height * (country==2)
gen IV1_dumFra_weight = IV1_weight * (country==2)
gen IV1_dumGer_horsepower = IV1_horsepower * (country==3)
gen IV1_dumGer_fuel = IV1_fuel * (country==3)
gen IV1_dumGer_width = IV1_width * (country==3)
gen IV1_dumGer_domestic = IV1_domestic * (country==3)
gen IV1_dumGer_height = IV1_height * (country==3)
gen IV1_dumGer_weight = IV1_weight * (country==3)
gen IV1_dumIta_horsepower = IV1_horsepower * (country==4)
gen IV1_dumIta_fuel = IV1_fuel * (country==4)
gen IV1_dumIta_width = IV1_width * (country==4)
gen IV1_dumIta_domestic = IV1_domestic * (country==4)
gen IV1_dumIta_height = IV1_height * (country==4)
gen IV1_dumIta_weight = IV1_weight * (country==4)
gen IV1_dumUK_horsepower = IV1_horsepower * (country==5)
gen IV1_dumUK_fuel = IV1_fuel * (country==5)
gen IV1_dumUK_width = IV1_width * (country==5)
gen IV1_dumUK_domestic = IV1_domestic * (country==5)
gen IV1_dumUK_height = IV1_height * (country==5)
gen IV1_dumUK_weight = IV1_weight * (country==5)
// Generating interactions of IV3 and country dummies
gen IV3_dumBel_horsepower = IV3_horsepower * (country==1)
gen IV3_dumBel_fuel = IV3_fuel * (country==1)
gen IV3_dumBel_width = IV3_width * (country==1)
gen IV3_dumBel_domestic = IV3_domestic * (country==1)
gen IV3_dumBel_height = IV3_height * (country==1)
gen IV3_dumBel_weight = IV3_weight * (country==1)
gen IV3_dumFra_horsepower = IV3_horsepower * (country==2)
gen IV3_dumFra_fuel = IV3_fuel * (country==2)
gen IV3_dumFra_width = IV3_width * (country==2)
gen IV3_dumFra_domestic = IV3_domestic * (country==2)
gen IV3_dumFra_height = IV3_height * (country==2)

```

```

gen IV3_dumFra_weight = IV3_weight * (country==2)
gen IV3_dumGer_horsepower = IV3_horsepower * (country==3)
gen IV3_dumGer_fuel = IV3_fuel * (country==3)
gen IV3_dumGer_width = IV3_width * (country==3)
gen IV3_dumGer_domestic = IV3_domestic * (country==3)
gen IV3_dumGer_height = IV3_height * (country==3)
gen IV3_dumGer_weight = IV3_weight * (country==3)
gen IV3_dumIta_horsepower = IV3_horsepower * (country==4)
gen IV3_dumIta_fuel = IV3_fuel * (country==4)
gen IV3_dumIta_width = IV3_width * (country==4)
gen IV3_dumIta_domestic = IV3_domestic * (country==4)
gen IV3_dumIta_height = IV3_height * (country==4)
gen IV3_dumIta_weight = IV3_weight * (country==4)
gen IV3_dumUK_horsepower = IV3_horsepower * (country==5)
gen IV3_dumUK_fuel = IV3_fuel * (country==5)
gen IV3_dumUK_width = IV3_width * (country==5)
gen IV3_dumUK_domestic = IV3_domestic * (country==5)
gen IV3_dumUK_height = IV3_height * (country==5)
gen IV3_dumUK_weight = IV3_weight * (country==5)
// IV estimator of model with heterogeneous price parameters
ivreg2 lsj_ls0 horsepower fuel width domestic height weight logpop loggdp i.country
i.year i.brand (price_dum* = IV1_dum* IV3_dum*), robust
// Testing the null hypothesis of homogeneous price coefficients
test price_dumBel = price_dumFra = price_dumGer = price_dumIta = price_dumUK
// Similar test but excluding UK
test price_dumBel = price_dumFra = price_dumGer = price_dumIta
// -----
// Question 2.1(f) Calculating elasticities & average elasticity
// -----
gen elasticity = (1-share)*(price_dumBel*_b[price_dumBel] + price_dumFra*_b[price_dumFra]
+ price_dumGer*_b[price_dumGer] + price_dumIta*_b[price_dumIta] + price_dumUK*_b[price_dumUK])
sum elasticity , detail
// -----
// Question 2.2(a) Calculating realized marginal cost
// -----
gen mcost = .

```

```

replace mcost = price + 1/(_b[price_dumBel]*(1-share)) if country==1
replace mcost = price + 1/(_b[price_dumFra]*(1-share)) if country==2
replace mcost = price + 1/(_b[price_dumGer]*(1-share)) if country==3
replace mcost = price + 1/(_b[price_dumIta]*(1-share)) if country==4
replace mcost = price + 1/(_b[price_dumUK] *(1-share)) if country==5
sum mcost, detail
// -----
//      Question 2.2(b) Fixed Effects estimation of Marginal Cost function
// -----
reg mcost horsepower fuel width domestic height weight i.country i.year i.brand,
vce(robust)
// -----
//      Question 2.2(c) IV estimation of Marginal Cost function including qu
// -----
ivreg2 mcost horsepower fuel width domestic height weight i.country i.year
i.brand (qu = IV1_dum* IV3_dum*), robust
log close

```