MSMA Program

Project 1

Pricing Analytics

Yuying He

Zheng Lian

Yingluo You

Tianyao Liu

Jing Zou

Question 4. Control Variables

4.1 Interpreting a Log-log Regression

After running the colleague's code, we found that the coefficient of log(eurpr) here equals to -0.296. This means that when price in european currency increases by 1%, the quantity sold would decrease by 0.296%. The intercept is 11.32, which means when the price is 1 Euro(log(0)), the log(qu) would be 11.322. The price elasticity is denoted by the coefficient of log(eurpr)(-0.296), which means the demand is inelastic. The absolute value is smaller than 1 and reflects that the consumer is not price sensitive.

However, the elasticity may not be reasonable because the demand in the automobile market is expected to be elastic (a higher |the coefficient of log(eurpr)|) in the short-term, though tend to be less elastic or inelastic in the long-term. That is because cars are durable goods and consumers could always delay their purchase with so many substitutes. Once they have bought a car, they do not need to buy in the near future, so in the long-term the demand should be inelastic.

4.2 Adding Control Variables

Variables we chose: We decided to add **li** as X, add **ye, co** and **ma** as fixed effects.
Reasons of why we chose or not chose these variables:
li / hp/ sp/ ac/ cy convey similar information. A higher value of these four factors, such as higher power and better cylinder volume of the car, reflects the car is less likely to achieve fuel efficiency. Consequently, those four variables have negative correlation with fuel efficiency, thus positively correlates with **li** value, providing similar effects on affecting the coefficient of log(eurpr) as well. We experimented with above variables, and found out that we can use **li** to replace the rest of them since there is few differences on coefficient.

co/ zcode/ brd/ type/ brand/ model/ cla/ home/ frm: There is no difference between using model or brand code in terms of their impacts on price's causal relationship related with Q. Other similar factors are **cla, home** and **frm**. So, we only used factor **co** as one of the fixed effects to better capture the relationship between car model and P or Q. Besides, **ye** variable as time fixed effects should be included in the regression undoubtedly and it has a strong relationship with **co** because car model in each time period is quite different and thus will have diverse impact on both P and Q. So we interacted **co** and **ye** to get better estimates.

pl/ do/ le/ wi/ he are unreliable or unrelated variables, thus we didn't use them.

ma/gdp: although GDP would affect demand quantity in a certain market, but the information that GDP conveys can also be delivered by ma(market). We have conducted a regression analysis using summary(lm(engdp~factor(ma),data=cardata) to verify these two variables relationship. The P value equals to 2.2e-16, so H0 can be rejected. This means that there is a significant relationship between engdp and factor(ma). Thus, we chose to drop gdp and just added ma into regression model as fixed effects variables.

Code:

```
reg4_3=felm(log(qu)~log(eurpr)+li|factor(ye):factor(co)+factor(ma),data=cardata)
summary(reg4_3)
```

The coefficient of log(eurpr) changes from -0.296 to -1.81051 and the coefficient of li is -0.20223.

5. Instrumental Variables

With the aim of investigating the change of quantity, only adding fixed effects may not be enough. We need to include instrumental variables, which only influence price from the cost side while do not influence the consumer demand.

Firstly, from the cost (input price) perspective, we assumed that the amount of iron ore used in each car model is proportional to the weight of that car. We created a variable named 'mac' by multiplying 'weight' and unit to represent the material cost. Next, we merged these two tables by 'year' and renamed it as 'cardata1'.

Secondly, from the institutional regulatory changes perspective, different country (market) has different tax regulations. We included 'tax' as another instrumental variable because even though the cost of the car are the same, the final price will be affected by the tax rate at different country. Also, tax in the middle of the supply chain (paid by the car manufacture and stores) does not influence consumer demand.

Code:

```
reg5=felm(log(qu)~li|factor(ye):factor(co)+factor(ma)|(log(eurpr)~tax+mac),data=cardata1)
summary(reg5)
```

The coefficient of log(eurpr) changes from -1.81051 to -3.22894 and the coefficient of li changes from -0.20223 to -0.14627 after adding the two instrumental variables. It means that after adding IVs, the customers of car would become more price sensitive. In other words, the price elasticity becomes more elastic. With 1% of the price increase, the demand of car will decrease by 3.23%.

6. Recovering Costs

We took subset of line 88 in 'cardata1', and renamed it as 'a' for next step analysis.
Now we know the optimal price set by the rival is 2350.145, and used the regression model we built in question 5 to estimate the demand. We used 'getfe' to estimate the total fixed effect imposed by the interaction of model code and year as will as the market fixed effect, renaming the sum of them as 'total_fe'.

Next, we used profit = (price-unit cost)*quantity to recovering costs. Below is our calculation process:

Firstly, calculate Q:

*log(Q) = the coefficient of log(eurpr)*log(price)+fe*

log(Q) = log[(price)^the coefficient of log(eurpr)]+log(fe)

*Q = P^the coefficient of log(eurpr) * e^fe*

Code:

```
Q = a$eurpr^reg4$beta[2] * exp(1)^(total_fe+a$li*reg4$beta[1])
```

Secondly, present profit. Then take derivative on the right side and make the left side as 0 to show the optimal position:

$$profit = P \ast P^{beta} \ast e^{(FE+CV)} - AVC \ast P^{beta} \ast e^{(FE+CV)}$$

$$profit = P^{(beta+1)} \ast e^{(FE+CV)} - AVC \ast P^{beta} \ast e^{(FE+CV)}$$

$$d(profit)/d(P) = reg4\$beta[2] \ast P - AVC \ast P^{(reg4\$beta[2]-1)} \ast e^{(FE+CV)} \ast reg4\$beta[2]$$

$$0 = P^{(beta)} \ast (beta+1) \ast e^{(FE+CV)} - AVC \ast P^{(beta-1)} \ast e^{(FE+CV)} \ast beta$$

$$\ast CV = li \ast \text{the coefficient of } log(eurpr)$$

Code:

```
0 = exp(1)^(total_fe++a$li*reg4$beta[1])*(reg4$beta[2]+1)*a$eurpr^reg4$beta[2]-reg4$beta[2]*a$eurpr^(reg4$beta[2]-1)*exp(1)^(total_fe+a$li*reg4$beta[1])* AVC
AVC1 = exp(1)^(total_fe+a$li*reg4$beta[1])*(reg4$beta[2]+1)*a$eurpr^reg4$beta[2]/reg4$beta[2]*a$eurpr^(reg4$beta[2]-1)*exp(1)^(total_fe+a$li*reg4$beta[1])
#AVC is 780.0417

AVC0 = exp(1)^(total_fe0+a$li*reg4$beta[1])*(reg4$beta[2]+1)*a$eurpr^reg4$beta[2]/reg4$beta[2]*a$eurpr^(reg4$beta[2]-1)*exp(1)^(total_fe0+a$li*reg4$beta[1])
```

By running this formula, we could get the rival's average production cost which is € 780.0417. As the rival's price is € 2350.1450, the cost takes up 33.2% of the revenue in the profit-maximization situation, while the unit contribution is € 1570.1033.

7. Cross-elastisties and Competitive Effects

The coefficient of the log(avgurprrival) is 5.72557 (cross-price elasticity), which represents that if average rival's price increases by 1%, the quantity of our car product sold would increase by 5.73%. This coefficient is cocsistent with our expectation, which is a positive value. The reason is that our rival's cars are substitutes of our car products, and if average rival's price increase, the demand for rivals' products would decrease and thus the demand for our car products would increase.

The coefficient of the log(eurpr) decreases from -3.22894 in Q5 to -5.44682, which indicates that there exists intensive competition in market because after incorporating rivals' price into regression analysis, the price elasticity of our car products demand is magnified. This means when substitutes exist, consumer would become more price sensitive toward our product because they could turn to our competitor if we price cars at a high level.

Code:

```
reg7_1=felm(log(qu)~li+log(avgurprrival)|factor(ye):factor(co)+factor(ma)|(log(eurpr)~tax+mac),data=cardata1)
summary(reg7_1)
```

Summary

In this report, our group first interpreted a Log-log Regression. In later sections, we built the regression model step by step by adding fixed effects, interactions, instrumental variables, and competitor's price. We also recovered competitor's cost by its optimal price and sales quantity.