|  |
| --- |
| Group Project 1: Estimating Coffee Demand and Supply Functions |
| Team Members: Yang Rui A0160930M, Bao Rui A0160941J, Chen Xiang A0160886U, Lyu Pei Jin A0161011E, Yau Chung Yin A0160936BDSC5101 ANALYTICS IN MANAGERIAL ECONOMICSSeptember 6, 2016 |

# Executive Summary

The paper seeks to estimate the coffee demand and supply functions and estimate the significant factors concerning the demand and supply function by using consumption and production data from the Dutch coffee market. We estimate the linear and logarithmic model by assuming the roasted coffee price is at market equilibrium. Our results indicate that mainly the bean price and income explain the movement of the roasted coffee price. We observed that tea is not viewed as a close substitute good of coffee, and suppliers tend to absorb the increment of other costs to stabilize the market price, thus making other costs less correlated with the market price.

# Introduction

In the past few years, price of coffee beans have experienced large fluctuations, whereas consumer price for roasted coffee hardly responded to the movement of coffee bean price. Hence, there is incomplete transmission of coffee bean price to consumer price, and it is affected by other factors.

The main objective of this paper is to estimate a structural model of coffee supply and demand function. With this objective, we use a time-series aggregate dataset that contains monthly information during the time period 1990-1996 on coffee consumption and production from the Dutch market.

This paper also seeks to evaluate the relative importance of cost and the significant and relational factors concerning the demand and supply function of the Dutch coffee market.

The structure of the paper is as follows. The next section will present a theoretical background for the empirical model and illustrate the influential factors relative to supply and demand function respectively. Section 3 discusses the formation of our model, estimation of the conduct parameter and determining the most reasonable factors under the circumstance of both log-linear regression and linear regression for supply and demand function. Section 4 reports the estimated parameters and covers the possible explanations. The last section summarizes the findings and draws conclusions.

# Model Assumptions

Our model estimates the coffee demand and supply function with an assumption that coffee is a homogeneous good because of the aggregate data provided.

From the economic perspective, the following factors affect the coffee demand. First and the most important factor is the price of coffee. An increase in the coffee price would lead to a decrease in consumption and vice versa. Second, the income level of consumers correlates with their buying behavior. An increase in the income of consumers could have a positive effect on the demand of goods in the market. Third, other available alternatives, such as tea, would be more favored if the price of tea is relatively more attractive. Finally yet importantly, the seasonality effect would influence all of the above factors when we specify the functional form. By considering all of the above factors, our demand function is as below:

QD=Q (, , )

On the other hand, supply largely depends on the cost of production. For the given data, we believe that the price of coffee beans and the labor cost are the most influencing factors for the supply function. Both of the cost factors negatively correlate to the quantity supply of coffee. Again, we include seasonality effect into our model. Our supply function is as follow by consolidating the above factors:

QS=Q (, , )

The cost factors listed above are not exhaustive. Other cost factors such as the cost of packaging and producing machine would also influence the supply of coffee, but these factors are not included in the data thus not included in the model.

# The Model

When using OLS to estimate the demand and supply function, we adopt the linear and logarithmic specification for both demand and supply equations:

(1)

(2)

(3)

(4)

Equation (1) and (2) represent the specifications of the demand side; Equation (3) and (4) are the supply equations.

Qu represents per capita consumption of roasted coffee in kg; cprice is the price of roasted coffee per kg in current guilders; income is income per capita in current guilders; tprice is the price of per kg tea in current guilders; bprice is price of coffee beans per kg in current guilders; wprice is the price of labor per man hours (work 160 hours per month); qi is the dummy variable for the ith quarter; ε is the error term. All the price or income variables are divided by price index for other goods (*oprice*).

From an empirical point of view, we assume that cprice is the equilibrium outcome of the interplay between the supply and demand. In this case, the cprice responds endogenously to demand shocks depending on the slope of the supply curves. Therefore, we treat cprice as an endogenous variable due to its correlations with the error term ε, and we begin the modeling with the TSLS (Two stage least squares) instead of OLS.

To start with, we first identify the control variables and the instrumental variables with the model we established in the previous session. We categorize variables as control or instrumental according to the significance level in the TSLS regression. In this regard for our linear model, we choose bprice as the instrumental variable on the demand side and income as the instrumental variable for the supply equation\*. We drop the variables tprice and wprice on this basis from the model and the equations are as follow:

(5)

(6)

Equation (5) is on the demand side and equation (6) is on the supply side. We use TSLS regression to first predict the cprice with control variables and instrument variables and then use the cprice predictions to regress with Qu. The values of Hausman test are lower than 0.05 in both Supply and Demand regressions. This suggests that our instrumental variables are unbiased and there is no over-identification.\*

The mechanism of performing TSLS regression of logarithmic specification is the same. We identify income and bprice are the instrumental variables on the demand and supply side respectively. The values of Hausman test are lower than 0.05 in both Supply and Demand regressions which shows the regression is statistically significant.\*

When comparing the linear and logarithmic model, the p-value in logarithmic model is relatively lower, which implies that the results are more significant. Furthermore, when it comes to the logarithmic form, the elasticity of the demand is more visible compared with the linear form. (The elasticity is β1, which is -0.278 in our case) We then adopt the logarithmic specification for our below discussion.

Also, we perform the OLS regression to see the differences in parameters. As expected, the differences exist due to the correlations between cprice and ε. All the parameters are in the table below.

Table 3.1 Parameters in linear demand

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| LINEAR(D) | α0 | α1 | α2 | α3 | α4 | α5 |
| TSLS | 0.5587 | -0.0137 | 0.0002 | -0.0792 | -0.0687 | -0.0760 |
| OLS | 0.5620 | -0.0130 | 0.0002 | -0.0794 | -0.0685 | -0.0762 |

Table 3.2 Parameters in logarithmic demand

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| LOG(D) | α0 | α1 | α2 | α3 | α4 | α5 |
| TSLS | -3.2738 | -0.2781 | 0.4898 | -0.1115 | -0.0925 | -0.1071 |
| OLS | -3.1300 | -0.2575 | 0.4636 | -0.1120 | -0.0920 | -0.1076 |

Table 3.3 Parameters in linear supply

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| LINEAR(S) | β0 | β1 | β2 | β3 | β4 | β5 |
| TSLS | 0.3852 | 0.0676 | -0.1399 | -0.0806 | -0.0606 | -0.0869 |
| OLS | 0.7621 | 0.0080 | -0.0334 | -0.0897 | -0.0676 | -0.0850 |

Table 3.4 Parameters in logarithmic supply

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| LOG(S) | β0 | β1 | β2 | β3 | β4 | β5 |
| TSLA | -2.1902 | 1.0606 | -0.6474 | -0.1118 | -0.0726 | -0.1137 |
| OLS | -0.3364 | 0.0899 | -0.1567 | -0.1258 | -0.0888 | -0.1567 |

Result discussion

Based on the analysis above, in logarithmic specification, we obtain the equations below:

1

*2*

When the market reaches its equilibrium, log(Qu)\_d = log(Qu)\_s, then we have:

From above equation, we observe that coffee price has a relationship with income and bprice. However, this result is totally different from our assumption, that price should be correlated with income, bprice, wprice, quarter and tprice variables. An explanation to strong relationship between income, bprice and cprice is that when customers have a higher income level, they are willing to pay for a higher price of coffee. For the suppliers they have to raise the price to cover the cost of coffee beans.

However, the situation is different for wprice and tprice. For wprice, one possible explanation is that the labor cost in the coffee industry accounts for a relatively small part of the total production cost. In case of increase in labor cost, the company may choose to absorb the increment and maintain a stable cprice to maximize their profits. In contrast, the high coefficient of bprice indicates its connection with cprice is significant in our regression. We believe the cost of coffee bean accounts for a relatively large proportion of total production cost. Therefore suppliers have to increase the cprice to avoid a great erosion in the marginal profit when bean price goes up.

For tprice factor, we think coffee drinkers, especially those with heavy consumptions, are not sensitive to movement of tea price. Even price of tea drops significantly, coffee customers are less likely to switch to tea. One explanation is consumers have different preference on drinking coffee over other drinks. In this case, tea may not be a close substitute of coffee. Therefore we believe that the movement of tprice will not cause a demand shock to the coffee market.

As we discussed in the second session of our paper, the dummy variable - quarter, will impact on the supply and demand curve. The result of our regression shows that are all negative (In our regression, q4 is our baseline as a dummy variable). That means the demand in q4 is higher than the demand in any other quarters. One explanation is that suppliers will need to prepare more roasted coffee to meet the demand for the festive season towards the end of the year.

# Conclusion

This paper has estimated the overall relatively significant and relational factors concerning the demand and supply function of the Dutch coffee market over the period from 1990 to 1996. By modeling the demand and supply function based on the original empirical model, we estimated the supply and demand functions.

The results indicate that the quantity in demand function is mainly affected by the income per capita in current guilders and the coffee consumer price. While the bean price and consumer price are the major influencing factors in the supply function, the significance of controlling for seasonality cannot be overstated both in the demand function and supply function. The tea price and labor price are omitted from empirical model as we consider these factors are relatively less significant to the demand and supply function.

On a final note, there is room for extensions of the model and further analysis. For example, we analyzed the demand and supply function from the perspective of the coffee industry as a whole. Further analysis can be conducted to acquire more comprehensive insights if firm-level dataset is available. Also, more quantitative data such as inventory cost, packaging cost, variable machine cost, and qualitative data including consumer loyalty, market structure would enable a more stringent model and resounding analysis on relationships between bean prices and consumer prices.

# Appendix

R code and detail outcome :

eco\_final.R

Tue Sep 06 09:22:01 2016

library(AER)

## Loading required package: car

## Loading required package: lmtest

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

## Loading required package: sandwich

## Loading required package: survival

library(MASS)  
#data preparation  
dataall<-read.csv("C:/Users/ACER/Desktop/MSBA/Economics/Projects/Project1/Project1Data.csv")  
dataall$bprice<-dataall$bprice/dataall$oprice  
dataall$wprice<-dataall$wprice/dataall$oprice  
dataall$tprice<-dataall$tprice/dataall$oprice  
dataall$incom<-dataall$incom/dataall$oprice  
dataall$cprice<-dataall$cprice/dataall$oprice  
##########################################  
#TSLS#linear part  
##########################################  
#determine the iv  
demand.linear<-lm(cprice~bprice+wprice+q1+q2+q3,data=dataall)  
summary(demand.linear)

##   
## Call:  
## lm(formula = cprice ~ bprice + wprice + q1 + q2 + q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.17975 -0.21594 -0.02715 0.19118 1.17531   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.1755017 4.0750972 -1.025 0.3087   
## bprice 1.8125348 0.0520762 34.805 <2e-16 \*\*\*  
## wprice 0.3571746 0.1384457 2.580 0.0118 \*   
## q1 -0.1643968 0.1438507 -1.143 0.2566   
## q2 -0.1671898 0.1451420 -1.152 0.2529   
## q3 -0.0001087 0.1441018 -0.001 0.9994   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.465 on 78 degrees of freedom  
## Multiple R-squared: 0.9412, Adjusted R-squared: 0.9374   
## F-statistic: 249.5 on 5 and 78 DF, p-value: < 2.2e-16

#demand function for linear 2sls  
demand.linear.stage1 <- lm(cprice~bprice+incom+q1+q2+q3, data=dataall)  
demand.linear.iv<-ivreg(qu~q1+q2+q3+incom+cprice|incom+bprice+q1+q2+q3,data=dataall)  
summary(demand.linear.iv)

##   
## Call:  
## ivreg(formula = qu ~ q1 + q2 + q3 + incom + cprice | incom +   
## bprice + q1 + q2 + q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10980 -0.04802 -0.01128 0.03786 0.28995   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.5586701 0.1763240 3.168 0.002189 \*\*   
## q1 -0.0792121 0.0227557 -3.481 0.000821 \*\*\*  
## q2 -0.0687351 0.0216801 -3.170 0.002176 \*\*   
## q3 -0.0759836 0.0220445 -3.447 0.000917 \*\*\*  
## incom 0.0001979 0.0001059 1.868 0.065561 .   
## cprice -0.0137198 0.0047472 -2.890 0.004985 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.06993 on 78 degrees of freedom  
## Multiple R-Squared: 0.2752, Adjusted R-squared: 0.2287   
## Wald test: 5.993 on 5 and 78 DF, p-value: 9.637e-05

#hausman test  
HT1 <- lm(formula = residuals(demand.linear.iv)~ incom, data = dataall)  
summary(HT1)

##   
## Call:  
## lm(formula = residuals(demand.linear.iv) ~ incom, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.10980 -0.04802 -0.01128 0.03786 0.28995   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1.321e-14 1.572e-01 0 1  
## incom -6.927e-18 8.753e-05 0 1  
##   
## Residual standard error: 0.06821 on 82 degrees of freedom  
## Multiple R-squared: 7.7e-29, Adjusted R-squared: -0.0122   
## F-statistic: 6.314e-27 on 1 and 82 DF, p-value: 1

Rsquare1 <- summary(HT1)$r.squared  
HT1\_result <- Rsquare1\*(84-1-1)  
HT1\_result

## [1] 6.313684e-27

#determine the iv  
supply.linear<-lm(formula=cprice~incom+tprice+q1+q2+q3,data=dataall)  
summary(supply.linear)

##   
## Call:  
## lm(formula = cprice ~ incom + tprice + q1 + q2 + q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.7574 -1.0568 -0.4954 0.4821 4.6033   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.715238 13.930484 -0.267 0.7904   
## incom 0.009525 0.003379 2.819 0.0061 \*\*  
## tprice -0.033807 0.501586 -0.067 0.9464   
## q1 0.253472 0.570140 0.445 0.6579   
## q2 -0.348591 0.533996 -0.653 0.5158   
## q3 0.319594 0.553007 0.578 0.5650   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.727 on 78 degrees of freedom  
## Multiple R-squared: 0.189, Adjusted R-squared: 0.137   
## F-statistic: 3.635 on 5 and 78 DF, p-value: 0.005225

##supply functioin for linear 2sls  
supply.linear.stage1 <- lm(cprice~bprice+incom+q1+q2+q3, data=dataall)  
supply.linear.iv<-ivreg(qu~q1+q2+q3+bprice+cprice|incom+q1+q2+q3+bprice,data=dataall)  
summary(supply.linear.iv)

##   
## Call:  
## ivreg(formula = qu ~ q1 + q2 + q3 + bprice + cprice | incom +   
## q1 + q2 + q3 + bprice, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.142611 -0.051618 -0.005987 0.048844 0.299560   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.38524 0.29072 1.325 0.189000   
## q1 -0.08057 0.02466 -3.267 0.001617 \*\*   
## q2 -0.06060 0.02427 -2.497 0.014633 \*   
## q3 -0.08692 0.02366 -3.674 0.000437 \*\*\*  
## bprice -0.13989 0.08193 -1.707 0.091721 .   
## cprice 0.06763 0.04563 1.482 0.142340   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.0765 on 78 degrees of freedom  
## Multiple R-Squared: 0.1327, Adjusted R-squared: 0.0771   
## Wald test: 5.008 on 5 and 78 DF, p-value: 0.0004963

###hausman test  
HT3<- lm(formula = residuals(supply.linear.iv)~incom, data = dataall)  
summary(HT3)

##   
## Call:  
## lm(formula = residuals(supply.linear.iv) ~ incom, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.142611 -0.051618 -0.005987 0.048844 0.299560   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1.534e-14 1.719e-01 0 1  
## incom -8.379e-18 9.574e-05 0 1  
##   
## Residual standard error: 0.07461 on 82 degrees of freedom  
## Multiple R-squared: 9.281e-29, Adjusted R-squared: -0.0122   
## F-statistic: 7.61e-27 on 1 and 82 DF, p-value: 1

Rsquare3<- summary(HT3)$r.squared  
HT3\_result <- Rsquare3\*(84-2)  
HT3\_result

## [1] 7.61017e-27

##########################################  
#TSLS#log part  
##########################################  
#determine the iv  
demand.log<-lm(formula=log(cprice)~log(bprice)+log(wprice)+q1+q2+q3,data=dataall)  
summary(demand.log)

##   
## Call:  
## lm(formula = log(cprice) ~ log(bprice) + log(wprice) + q1 + q2 +   
## q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.095309 -0.021506 -0.003358 0.023447 0.090818   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.22171 1.18361 -1.877 0.064249 .   
## log(bprice) 0.52171 0.01780 29.311 < 2e-16 \*\*\*  
## log(wprice) 1.21937 0.34926 3.491 0.000794 \*\*\*  
## q1 -0.01567 0.01225 -1.279 0.204652   
## q2 -0.02261 0.01236 -1.829 0.071240 .   
## q3 -0.00796 0.01229 -0.648 0.518997   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.03964 on 78 degrees of freedom  
## Multiple R-squared: 0.9194, Adjusted R-squared: 0.9142   
## F-statistic: 177.9 on 5 and 78 DF, p-value: < 2.2e-16

#demand function for log 2sls  
demand.linear.stage1 <- lm(log(cprice)~log(bprice)+log(incom)+q1+q2+q3, data=dataall)  
demand.log.iv<-ivreg(log(qu)~q1+q2+q3+log(incom)+log(cprice)|log(bprice)+q1+q2+q3+log(incom),data=dataall)  
summary(demand.log.iv)

##   
## Call:  
## ivreg(formula = log(qu) ~ q1 + q2 + q3 + log(incom) + log(cprice) |   
## log(bprice) + q1 + q2 + q3 + log(incom), data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.187313 -0.071738 -0.007799 0.062038 0.333603   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.27378 1.90013 -1.723 0.088865 .   
## q1 -0.11152 0.03189 -3.498 0.000778 \*\*\*  
## q2 -0.09246 0.03033 -3.049 0.003137 \*\*   
## q3 -0.10709 0.03086 -3.470 0.000850 \*\*\*  
## log(incom) 0.48976 0.26559 1.844 0.068971 .   
## log(cprice) -0.27806 0.09308 -2.987 0.003761 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09786 on 78 degrees of freedom  
## Multiple R-Squared: 0.2782, Adjusted R-squared: 0.2319   
## Wald test: 6.109 on 5 and 78 DF, p-value: 7.967e-05

#hausman test  
HT2<- lm(formula = residuals(demand.log.iv) ~ log(bprice), data = dataall)  
summary(HT2)

##   
## Call:  
## lm(formula = residuals(demand.log.iv) ~ log(bprice), data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.18731 -0.07174 -0.00780 0.06204 0.33360   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -4.642e-15 5.340e-02 0 1  
## log(bprice) -4.202e-16 4.128e-02 0 1  
##   
## Residual standard error: 0.09545 on 82 degrees of freedom  
## Multiple R-squared: 1.384e-30, Adjusted R-squared: -0.0122   
## F-statistic: 1.135e-28 on 1 and 82 DF, p-value: 1

Rsquare2 <- summary(HT2)$r.squared  
HT2\_result <- Rsquare2\*(84-1-1)  
HT2\_result

## [1] 1.134865e-28

#determine the iv  
supply.log<-lm(formula=log(cprice)~log(tprice)+log(incom)+q1+q2+q3,data=dataall)  
summary(supply.log)

##   
## Call:  
## lm(formula = log(cprice) ~ log(tprice) + log(incom) + q1 + q2 +   
## q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.20093 -0.08229 -0.02724 0.04299 0.31365   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6.66536 4.78489 -1.393 0.16758   
## log(tprice) -0.05224 0.64592 -0.081 0.93574   
## log(incom) 1.24844 0.43680 2.858 0.00546 \*\*  
## q1 0.02105 0.04147 0.508 0.61321   
## q2 -0.02339 0.03879 -0.603 0.54831   
## q3 0.02488 0.04016 0.619 0.53746   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1254 on 78 degrees of freedom  
## Multiple R-squared: 0.1931, Adjusted R-squared: 0.1414   
## F-statistic: 3.733 on 5 and 78 DF, p-value: 0.004408

#supply function for log 2sls  
supply.linear.stage1 <- lm(log(cprice)~log(bprice)+log(incom)+q1+q2+q3, data=dataall)  
supply.log.iv<-ivreg(log(qu)~q1+q2+q3+log(bprice)+log(cprice)|log(bprice)+q1+q2+q3+log(incom),data=dataall)  
summary(supply.log.iv)

##   
## Call:  
## ivreg(formula = log(qu) ~ q1 + q2 + q3 + log(bprice) + log(cprice) |   
## log(bprice) + q1 + q2 + q3 + log(incom), data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.231753 -0.078798 -0.002549 0.076108 0.360204   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.19023 1.44461 -1.516 0.133525   
## q1 -0.11179 0.03490 -3.203 0.001970 \*\*   
## q2 -0.07257 0.03547 -2.046 0.044139 \*   
## q3 -0.11370 0.03324 -3.420 0.000999 \*\*\*  
## log(bprice) -0.64736 0.38478 -1.682 0.096491 .   
## log(cprice) 1.06060 0.75568 1.404 0.164432   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.1073 on 78 degrees of freedom  
## Multiple R-Squared: 0.1328, Adjusted R-squared: 0.07723   
## Wald test: 5.085 on 5 and 78 DF, p-value: 0.0004364

#hausman test  
HT4<- lm(formula = residuals(supply.log.iv)~log(incom), data = dataall)  
summary(HT4)

##   
## Call:  
## lm(formula = residuals(supply.log.iv) ~ log(incom), data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.23175 -0.07880 -0.00255 0.07611 0.36020   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 6.485e-13 1.788e+00 0 1  
## log(incom) -8.700e-14 2.387e-01 0 1  
##   
## Residual standard error: 0.1046 on 82 degrees of freedom  
## Multiple R-squared: 1.617e-27, Adjusted R-squared: -0.0122   
## F-statistic: 1.326e-25 on 1 and 82 DF, p-value: 1

Rsquare4 <- summary(HT4)$r.squared  
df4 <- HT4$df.residual  
HT4\_result <- Rsquare4\*(84-1-1)  
HT4\_result

## [1] 1.326024e-25

##########################################  
#OLS part  
##########################################  
#determine the demand function for linear  
demand.linear.ols<-lm(qu~cprice+incom+q1+q2+q3,data=dataall)  
summary(demand.linear.ols)

##   
## Call:  
## lm(formula = qu ~ cprice + incom + q1 + q2 + q3, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.11148 -0.04719 -0.01190 0.03848 0.29052   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.5620276 0.1762050 3.190 0.002052 \*\*   
## cprice -0.0129911 0.0045855 -2.833 0.005865 \*\*   
## incom 0.0001908 0.0001052 1.813 0.073715 .   
## q1 -0.0794018 0.0227498 -3.490 0.000797 \*\*\*  
## q2 -0.0684802 0.0216724 -3.160 0.002247 \*\*   
## q3 -0.0762216 0.0220373 -3.459 0.000882 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.06992 on 78 degrees of freedom  
## Multiple R-squared: 0.2754, Adjusted R-squared: 0.2289   
## F-statistic: 5.929 on 5 and 78 DF, p-value: 0.000107

###supply function for linear ols  
supply.linear.ols<-lm(formula = qu~cprice+q1+q2+q3+bprice,data=dataall)  
summary(supply.linear.ols)

##   
## Call:  
## lm(formula = qu ~ cprice + q1 + q2 + q3 + bprice, data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.114999 -0.044507 -0.009286 0.039886 0.291658   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.762082 0.109840 6.938 1.03e-09 \*\*\*  
## cprice 0.008024 0.016556 0.485 0.629271   
## q1 -0.089687 0.022049 -4.068 0.000113 \*\*\*  
## q2 -0.067618 0.022003 -3.073 0.002917 \*\*   
## q3 -0.085015 0.021876 -3.886 0.000212 \*\*\*  
## bprice -0.033432 0.030577 -1.093 0.277600   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.07084 on 78 degrees of freedom  
## Multiple R-squared: 0.2563, Adjusted R-squared: 0.2086   
## F-statistic: 5.375 on 5 and 78 DF, p-value: 0.000268

#demand function for log ols  
demand.log.ols<-lm(formula=log(qu)~log(cprice)+q1+q2+q3+log(incom),data = dataall)  
summary(demand.log.ols)

##   
## Call:  
## lm(formula = log(qu) ~ log(cprice) + q1 + q2 + q3 + log(incom),   
## data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.19074 -0.06991 -0.00865 0.05869 0.33469   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.13001 1.88835 -1.658 0.101428   
## log(cprice) -0.25754 0.08832 -2.916 0.004627 \*\*   
## q1 -0.11197 0.03187 -3.513 0.000739 \*\*\*  
## q2 -0.09198 0.03031 -3.035 0.003271 \*\*   
## q3 -0.10761 0.03084 -3.489 0.000799 \*\*\*  
## log(incom) 0.46362 0.26286 1.764 0.081690 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09783 on 78 degrees of freedom  
## Multiple R-squared: 0.2787, Adjusted R-squared: 0.2325   
## F-statistic: 6.028 on 5 and 78 DF, p-value: 9.099e-05

#supply function for log ols  
supply.log.ols<-lm(formula = log(qu)~log(cprice)+q1+q2+q3+log(bprice),data=dataall)  
summary(supply.log.ols)

##   
## Call:  
## lm(formula = log(qu) ~ log(cprice) + q1 + q2 + q3 + log(bprice),   
## data = dataall)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.19342 -0.06066 -0.01028 0.06009 0.33937   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.33640 0.50551 -0.665 0.507709   
## log(cprice) 0.08990 0.26287 0.342 0.733262   
## q1 -0.12583 0.03081 -4.084 0.000107 \*\*\*  
## q2 -0.08883 0.03089 -2.876 0.005190 \*\*   
## q3 -0.11760 0.03056 -3.848 0.000242 \*\*\*  
## log(bprice) -0.15671 0.13962 -1.122 0.265146   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.09897 on 78 degrees of freedom  
## Multiple R-squared: 0.2619, Adjusted R-squared: 0.2145   
## F-statistic: 5.534 on 5 and 78 DF, p-value: 0.0002056