Assignment 2: The second on multiple regression - Elasticity and nonlinear models

Barcelona Graduate School of Economics
Msc in Competition and Market Regulation

Course: Econometrics
Prof.: Javier Gómez Biscarri

Bendegúz István Biksi, Xia Mengzhen

November 5, 2015

1.

a.)

Compute the regression of per capita consumption of gasoline (G/pop) on all other explanatory variables, including the time trend (you can use either Year directly or (Year-1960)). Comment on whether the signs of the estimates agree with your expectations (i.e. comment on your expectations, too!).
The estimated model is:

$$G/pop_i = \beta_0 + \beta_1 * pg_i + \beta_2 * y_i + \beta_3 * pnc_i + \beta_4 * puc_i + \beta_5 * ppt_i + \beta_6 * pd_i + \beta_7 * pn_i + \beta_8 * ps_i + \beta_9 * year_i + \epsilon_i$$

The variables are:
G – total US gasoline consumption (total expenditure divided by price index)
Pg – price index for gasoline
Y – per capita disposable income
Pnc – price index for new cars
Puc – price index for used cars
Ppt – price index for public transportation
Pd – aggregate price index for consumer durables
Pn – aggregate price index for consumer nondurables
Ps – aggregate price index for consumer services
Pop – US total population in millions
Year – year of the data

Variables	g/pop
year	0.00940
	(0.00852)
pg	-0.121***
	(0.0253)
y	0.000111***
	(2.72e-05)
pnc	0.0635
	(0.112)
puc	-0.0408
	(0.0253)
ppt	0.0589
	(0.0369)
pd	0.290
	(0.370)
pn	0.542
	(0.420)
ps	-0.877***
	(0.304)
Constant	-18.43
	(16.51)
Observations	36
R-squared	0.983

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Before the regression, we could say that the price index for used and for new cars and price index for gasoline will have negative effect on the consumption of gasoline, since as for the consumer, cars and gasoline are complementary goods. So if the price of cars increases, the amount of gasoline consumed will decrease. Meanwhile, if the price of gasoline is higher, the unit consumption will decrease, because consumers has to cut off using it to save money. (Assuming gasoline is an ordinary good). We also think that, as the disposable income increases, the gasoline consumption increases too, so we assume, that gasoline is a normal good. The price index for gasoline have a significant negative coefficient, meanwhile the coefficients of the price indexes of cars are positive and negative, but not significant. As the income goes up by 1%, the percapita gasoline consumption will do so, but with a smaller percentage, than 1. Public transportation is assumed to be a substitute to gasoline, so ppt should have a positive coefficient, because as the price of the public transportation goes up, consumers substitute to using cars, therefore the consumed amount of gasoline would go up. Surprisingly the coefficient of ppt is not significant in 5% significance level. The coefficient of the time variable (year) is positive and significant. Meaning that, by time passes, gasoline consumption goes up, holding everything the same.

b.)

Test the hypothesis that consumers do not differentiate between changes in the prices of new and used cars.

$$\begin{aligned} H_0 : \beta_{pnc} - \beta_{puc} &= 0 \\ F(1, 26) &= 0.91 \\ Prob > F &= 0.3481 \end{aligned}$$

Since the P-value is bigger than 0.05, we can accept the null hypothesis, so consumers do not differentiate between changes in the prices of new and used cars.

c.)

Find the value of the elasticity of gasoline demand to its own price, to income and the cross-price elasticity with respect to changes in the price of public transportation from the results of this linear model (you will have to do a bit of thinking here, and remember the concept of elasticity).

Variables	pg	y	ppt
Elasticity	-0.278***	1.021***	0.161
	(0.058)	(0.250)	(0.101)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

First of all, we took a look at the elasticities at the means of the variables, that were quite similar to the avarage elasticities.

As for the price index of gasoline, it has negative elasticity on gas consumption. However, it's elasticity is smaller than one, which means that when the price index increased by 1%, the consumption of gas will decrease by a smaller extent than 1%.

The results above show that the income and public transportation has positive elasticity, that is to say when the income and public transportation changed, the price of gasoline will give responses at the same direction. Because the elasticity of ppt is smaller than 1, meaning inelasticity, it has slightly impacts on the gasoline consumption. However it turns out, that the elasticity of ppt is not significant. While the elasticity of per capita disposable income is higher than one, which means that when the per capita disposable income increased by 1%, the consumption of gas will increase to a larger extent than 1%.

d.)

Reestimate the above regression in logarithms (except for the time trend). All the coefficients are now directly elasticities. How do the estimates compare with the results in the previous analysis?
The estimated model is:

$$log(G/pop_i) = \beta_0 + \beta_1 * logpg_i + \beta_2 * logy_i + \beta_3 * logpnc_i + \beta_4 * logpuc_i + \beta_5 * logppt_i + \beta_6 * logpd_i + \beta_7 * logpn_i$$

$$+\beta_8 * logps_i + \beta_9 * year_i + \epsilon_i$$

The log-log model has other coefficients, regarding the elasticities. The own price elasticity changed by 92% as decreasing from $-.278$ to $-.535$. Nevertheless this log-log estimate still gives inelastic demand considering the own price elasticity. The elasticity of the gasoline percapita consumption with respect to the price index of the public transportation has also changed. Namely it has fallen by 25% from .161 to .121, however it is still insignificant. The income elasticity changed by 19% from 1.021 to 1.215. So this would imply that gasoline is a luxury good, since it's income elasticity is higher than 1.

Here the coefficient of the time variable (year) is negative, but insignificant. One is able to see that in the log-log model the number of the significant explanatory variables is higher compared to the linear model. On the other hand one should also take care using these models for forecasting because of the finite number of the sample size.

Variables	log(G/pop)
logPG	-0.535***
	(0.0596)
logY	1.215***
	(0.155)
logPPT	0.121
	(0.0712)
logPNC	0.0831
	(0.200)
logPUC	-0.115*
	(0.0634)
logPD	0.943***
	(0.254)
logPN	1.217***
	(0.237)
logPS	-1.307***
	(0.345)
year	-0.00440
	(0.00658)
Constant	-1.808
	(12.45)
Observations	36
R-squared	0.989

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

e.)

Do the two regression models imply the same regarding the elasticities? Which one do you think is more correct: the linear model or the log-log model? Why?

With the linear model we assume the same slope along the whole demand curve but variable elasticity. In the log-log model we assume constant elasticity, but varying slope. In the first model we took a look at the means of every variable to determine the elasticities. According to this we got the elasticities in one particular point (at the means of each variable). In the log-log model where the elasticities are the same along the whole demand curve we also determined the elasticities, but not just in one particular point (rather along the whole curve). With the constant elasticity we assume constant budget demand, meaning one should spend the same amount of money on the good regardless the price of the good. In the case of percapita gasoline consumption it is an acceptable assumption, that consumers pay the same amount of money and do not pay to much attention to price. So with this theoretical consideration and with the statistical results, namely more significant explanatory variables occur in the second model one should use the log-log model to scan the demand of gasoline.

2.

a.)

Use nonlinear least squares to get estimates of the elasticities of output to labor and capital. The estimated model is:

$$Y_i = A * L_i^{\beta} * K_i^{\gamma} + e_i$$

Where A is a composite factor ("technology"), L is labor and K is capital. The parameters β and γ measure the elasticity of output to labor and capital.

Variables	alpha	beta	gama
Constant	2.736***	0.404**	0.551***
	(0.921)	(0.167)	(0.133)
Observations	27		
R-squared	0.981		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The elasticity of output to labor and to capital are both significant in .05 significance level. The same is true for the technological parameter A. As the used amount of labor increases with 1%, the output increases by .404%. As the used amount of capital increases with 1%, the output increases by .551%. The R^2 is quite high. The sample size is not to big as it contains 27 observations.

b.)

Transform now this nonlinear model by taking logs of value added, labor and capital. Notice that the model now is perfectly linear, so it can be estimated by OLS. Do it, and compare the estimates of A, β and γ. Have these changed much?
The estimated model is:

$$logY_i = log(A * L_i^{\beta} * K_i^{\gamma} * e^{e_i})$$

That is the same as:

$$logY_i = logA + \beta * logL_i + \gamma * logK_i + e_i$$

Have the parameters changed much? Of course it depends on what do we call a large change. The elasticity of output to labor changed from .40 to .60 and the elsticity of output to capital changed from .55 to .38. The composite factor A changed from 2.74 to 1.17. So one could say they changed a lot, because the percentage change in the case of the labor is 50%, as in the case of the capital the percentage change is -32%. The composite factor changed by -57%. But the two elasticities are still ad up to .96. The change is structural. With the non-linear estimation Y was a capital intensive product, meanwhile with the OLS estimation Y seems a labor intensive product, because in the second case the elasticity of the labor is higher compared to the elasticity of the capital.

Variables	Log-Log with OLS
log(Labor)	0.603***
	(0.175)
log(Capital)	0.376***
	(0.115)
Constant	1.171***
	(0.281)
Observations	27
R-squared	0.943

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

c.)

Test the following hypotheses:

a. Labor elasticity is equal to 0.3.

$$\begin{aligned} H_0 : \beta &= 0.3 \\ F(1,24) &= 2.99 \\ Prob > F &= 0.0965 \end{aligned}$$

As the P-value of the F-statistic is higher, then .05 we can not reject the null hypothesis.

b. Constant returns to scale ($\beta + \gamma = 1$).

$$
\begin{aligned}
H_0 : \beta + \gamma &= 1 \\
F(1, 24) &= 0.09 \\
Prob > F &= 0.7608
\end{aligned}
$$

Here as the P-value of the F-statistic is higher, then .05 we can not reject the null hypothesis of constant return to scale.

c. Labor elasticity is equal to 0.3 and the production function has constant returns to scale.

$$
\begin{aligned}
H_0 : \beta &= 0.3 \\
H_0 : \beta + \gamma &= 1 \\
F(2, 24) &= 12.65 \\
Prob > F &= 0.0002
\end{aligned}
$$

As the P-value is smaller, then .05 we can reject the null hypothesis, so these to assumption do not hold ad the same time, however they hold separetly.

d.)

Test whether the translog function is more appropriate than the Cobb-Douglas function (so, in a sense, test the linearity of lny or that the Cobb-Douglas function is enough to account for output). Are the coefficients of the translog function as easy to interpret as those of the Cobb-Douglas function?

Ramsey RESET test using powers of the fitted values of Car Price
Ho: model has no omitted variables
F(3, 21) = 0.08
Prob > F = 0.9698

Here as the P-value is higher than .05 we can not reject the H_0, so the original Log-Log modell (the Cobb-Douglas production function) captures enough features, since the H_0 states, that there is no omitted variable. According to the test, the linearity of lnY in lnK and in lnL seems reasonable.

The coefficients of the translog function are not as easy to interpret as those of the Cobb-Douglas function. In addition to this it turns out that non of the parameters from the translog function estimate has a significant explanatory power. All in all it seems, that the Cobb-Douglas production function is more useful in this scenario.

Variables	$logY$
$logLabor$	3.614
	(2.158)
$logCapital$	-1.893
	(1.214)
$.5 * log^2 Labor$	-0.964
	(0.621)
$.5 * log^2 Capital$	0.0853
	(0.309)
$logLabor * logCapital$	0.312
	(0.369)
$Constant$	0.944
	(2.791)
Observations	27
R-squared	0.955

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1