

Quantitative Techniques – Term 2

Laboratory 4

9 February 2006

Overview

The objectives of this lab are to:

- Estimate demand elasticity;
- Show the effects on coefficients of a misspecified model/violation of OLS assumptions;
- Introduce ‘do’ files.

Start Stata 9 from the Windows Start menu (Programs, F. Departmental Software, E. Social Science, Stata). Do not update the version that is installed on your machine.

Before starting your tasks, make sure that the following windows are visible on your screen and choose their size so that they do not overlap: Review, Variables, Results, Command.

Task 1 – Price Elasticity Analysis

Go to www.staff.city.ac.uk/a.banal-estanol/teaching.htm and download Lab3Demand.dta to your directory.¹ Open it in Stata.

The file contains annual data on the U.S. gasoline market for the years 1960-1995. The variables are (in natural logarithms):

- Lngpop: U.S. gasoline consumption per capita (deflated);
- Lnpng: price of gasoline;
- Lnyc: per capita income;
- Lnpnc: price of new cars;
- Lnpus: price of used cars;

We will use this dataset to estimate own-price elasticities. Reminder: “The own-price elasticity of a product (or group of products) is defined as the percentage change in sales for each percentage change in the price.” (Lexecon report, page 14)

¹ The source is: Greene (2003), op. cit, Example 2.3., page 12.

1. OLS regression²

Under the OLS assumptions, we estimate the following model:

$$\text{Lngpop} = \beta_1 + \beta_2 * \text{lnpg} + \beta_3 * \text{lny} + \beta_4 * \text{lnpnc} + \beta_5 * \text{lnpuc} + \beta_6 * \text{year} + \text{error}$$

```
. reg lngpop lnpg lny lnpnc lnpuc year
```

Source	SS	df	MS			
Model	.780166138	5	.156033228	Number of obs =	36	
Residual	.025187728	30	.000839591	F(5, 30) =	185.84	
Total	.805353866	35	.02301011	Prob > F =	0.0000	
				R-squared =	0.9687	
				Adj R-squared =	0.9635	
				Root MSE =	.02898	

lnpop	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpg	-.1155303	.0334795	-3.45	0.002	-.1839046	-.0471561
lny	1.954627	.1928542	10.14	0.000	1.560766	2.348488
lnpnc	.2052824	.1520193	1.35	0.187	-.1051825	.5157473
lnpuc	-.1292738	.0714118	-1.81	0.080	-.2751161	.0165684
year	-.0191182	.0059565	-3.21	0.003	-.031283	-.0069533
_cons	20.06665	10.11464	1.98	0.056	-.5901884	40.7235

Given that the model is in loglinear form (or constant elasticity form), we can read the elasticities of demand with respect to the regressors directly from the coefficients.

As expected, the own-price elasticity is negative and pretty small in absolute value (-0.11), which indicates that gasoline demand is not very elastic. The income elasticity also has the expected sign.

2. Omitted variables/Misspecification of the model

Note what would happen if income was omitted from the above equation, i.e. if we were to estimate the following model.

```
. reg lngpop lnpg lnpnc lnpuc
```

Source	SS	df	MS			
Model	.411551052	3	.137183684	Number of obs =	36	
Residual	.393802877	32	.01230634	F(3, 32) =	11.15	
Total	.80535393	35	.023010112	Prob > F =	0.0000	
				R-squared =	0.5110	
				Adj R-squared =	0.4652	
				Root MSE =	.11093	

lnpop	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpg	.0756982	.1061912	0.71	0.481	-.1406062	.2920026
lnpnc	-.6680671	.4145029	-1.61	0.117	-1.512382	.1762478
lnpuc	.4948794	.2484313	1.99	0.055	-.0111585	1.000917
_cons	-.0870583	.033138	-2.63	0.013	-.1545582	-.0195584

We obtain the counterintuitive result that the own-price elasticity is positive (+0.07), and the coefficient is not significant. The coefficient on Lnpng is in fact capturing also the

² Note that the constant reported in Greene (2003), Example 4.4, should be corrected: see <http://pages.stern.nyu.edu/~wgreene/Text/Errata/ERRATA5.htm>

relationship between price and per capita income (the omitted variable). E.g. see Lecture 3, or Ashelfelter et al., or Greene (2003), Section 8.2., for a technical explanation.

3. Violations of OLS assumptions/Misspecification: serially correlated residuals³

So far we have estimated the demand model under the OLS assumptions, but we have not verified whether this was correct. Tests on the residuals from the OLS estimation should have been carried out to detect potential problem. For instance, a misspecification of the model may also result in problems of serial correlation in the residuals.

Assume we regress gasoline consumption only on its price by OLS, i.e.

```
. reg lngpop lnpg
```

Source	SS	df	MS	Number of obs = 36		
Model	.359241235	1	.359241235	F(1, 34)	=	27.38
Residual	.446112631	34	.01312096	Prob > F	=	0.0000
				R-squared	=	0.4461
				Adj R-squared	=	0.4298
				Root MSE	=	.11455

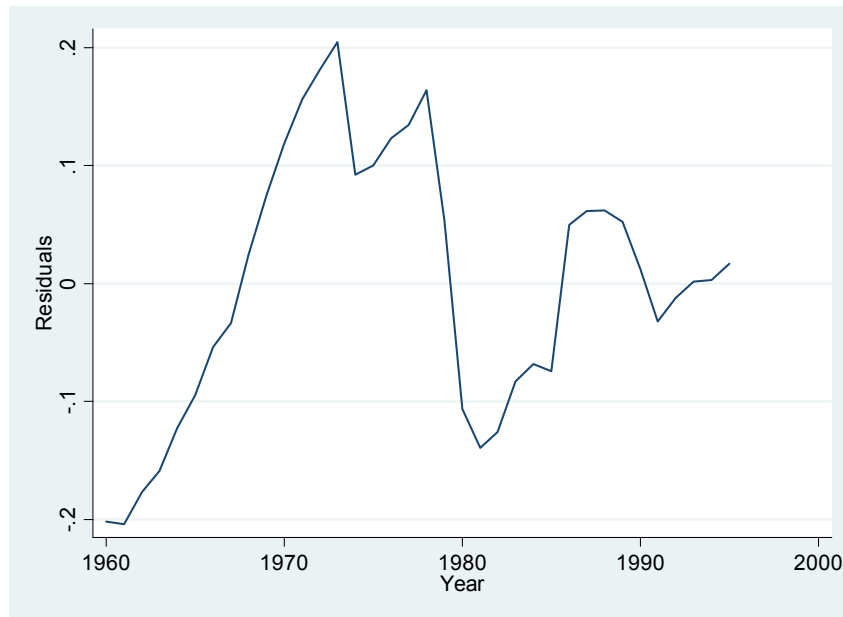
lnpgop	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnpg	.1676711	.0320441	5.23	0.000	.1025497	.2327925
_cons	-.1167347	.0288282	-4.05	0.000	-.1753206	-.0581489

In the model above, the price elasticity is again positive, as in point 2 above.

Note that the residuals from this misspecified equation would look serially correlated:⁴

³ This point is based on example 12.2. of Greene (2003).

⁴ Reminder from Lab1 – After you estimate a model, you can obtain the residuals by typing: predict resid, residuals. Note that ‘resid’ is the name I give to the new variable that contains the residuals, but you could choose another one.

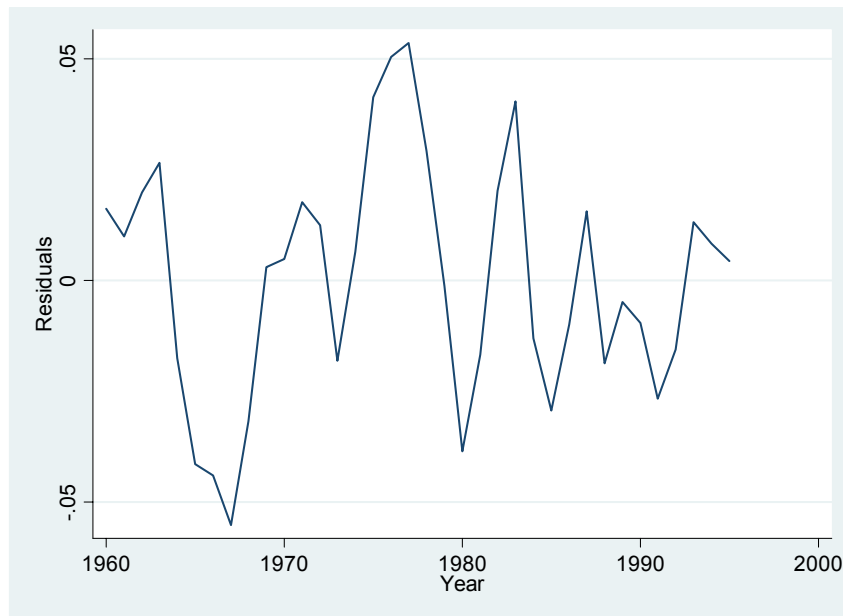


You can obtain the plot in two ways:

- In the Command window, type: `line resid year`
- From the Menu, choose Graphics... Easy Graph... Line graph.

If we compare this with the residuals from the full model, as in point 1, we can see that improving the specification of the model reduces the autocorrelation of the residuals.

The residuals from the first model have already been created in the dataset and are called 'resfull'. Plot them using the commands above.



By improving the specification of the model, the degree of serial correlation in the residuals has been reduced and the sign of the own-price elasticity has become negative.

Summary of Task 1:

- The estimation of demand elasticities is carried out in the context of a model of consumption as a function of income, own price and other factors;
- In our example, we found a low elasticity of demand with respect to price. A hypothetical monopolist would therefore find it profitable to increase prices, as the reduction in sales would be outweighed by the higher profits;
- A misspecified model may bias coefficients (example of omitted variables);
- With time series data, serial correlation of the residuals may arise. We have seen an example in which it can be caused by a misspecification of the model.

Given the results in point 1, what can you conclude about the definition of the market for gasoline?

What other example of violation of the OLS assumptions can you think of?

Task 2 – Create a ‘do’ file⁵

Stata can be used in interactive mode (what we have done so far: one command at a time from the Command window) or you can write the sequence of commands you need in a separate file and Stata will execute them all. This separate file is called a ‘do’ file.

This is particularly useful if you are going to write many commands and you want to be able to reproduce your work later without typing all commands again.

- In order to **open** the Stata ‘do’ file editor, you can:
 - On the Toolbar, click on the icon which looks like an envelope; or
 - On the Menu bar, click on Windows... Do-file editor.
- Now you have a blank page in which you can **type** the commands we have used in this lab. Press the ‘Enter’ key after each command.
- **Save** your do file in your directory.
- In the do-file editor, **execute the commands** in the do file by clicking Tools... Do on the Menu bar. You can also ‘do’ only some selected commands in your file: select the commands and then choose ‘Do selection’ from Tools.

1. Tips

- You can open (and modify) a do file from a text editor, e.g. NotePad

⁵ Definition (www.stata.com): A ‘do’ file is a plain text file with no special characters containing Stata commands that you create with an editor or word processor.

- You can insert comments in the file (i.e. text which Stata does not read as a command) between `/*` and `*/`. For instance, Stata ignores the following statement `/* this is a comment */`
- In the results window, *-more-* may appear when you are trying to output a long listing. Then you need to click on it for the listing to continue in the results window. The command `'set more off'` these interruptions and it is useful at the beginning of a do file.

2. Log files

All output appearing in the results window can be captured in a log file. The log file can be saved as a Stata formatted or as a text file.

- To start a log AS A TEXT:
`log using "C:\your directory\filename.log"` [starts a text log]
- To close a log:
`log close` [closes current log file]
- Remember to specify the directory and the file name within inverted commas.
- You can include the above commands in your do file to save your output in a text file (for example, see Demand.do).