

*For the class in the week beginning January 16*

**PROBLEM SET 10**

**(0)** Write down the name and office hour of your class teacher.

**(1)\*** Exercise 7.3 in the text

Fit an earnings function using your *EAEF* data set, taking *EARNINGS* as the dependent variable and *S*, *EXP*, and *MALE* as the explanatory variables, and perform a Goldfeld–Quandt test for heteroscedasticity in the *S* dimension. Remember to sort the observations by *S* first.

*Stata commands*

```
sort S
reg EARNINGS S EXP MALE in 1/203
reg EARNINGS S EXP MALE in 338/540
```

**(2)\*** Exercise 7.4 in the text

Fit an earnings function using your *EAEF* data set, using the same specification as in Exercise 7.3 and perform a White test for heteroscedasticity. Explain why Stata drops *MALESQ* from the regression.

*Stata commands*

```
reg EARNINGS S EXP MALE
predict EEARN, resid
gen EEARNsq = EEARN^2
gen SSQ = S^2
gen EXPSQ = EXP^2
gen MALESQ = MALE^2
gen MALES = MALE*S
gen MALEEXP = MALE*EXP
gen SEXP = S*EXP
reg EEARNsq S EXP MALE SSQ EXPSQ MALESQ MALES MALEEXP SEXP
```

**(3)\*** Exercise 7.8 in the text

Using your *EAEF* data set, repeat Exercises 7.3 and 7.4 with *LGEARN* as the dependent variable. Is there evidence that this is a preferable specification?

*Stata commands*

```
sort S
reg LGEARN S EXP MALE in 1/203
reg LGEARN S EXP MALE in 338/540
reg LGEARN S EXP MALE
predict ELGEARN, resid
gen ELGSQ = ELGEARN^2
reg ELGSQ S EXP MALE SSQ EXPSQ MALESQ MALES MALEEXP SEXP
```

For the class in the week beginning January 23:

### PROBLEM SET 11

**(1)** Exercise 8.6 in the text

Suppose that the true model is  $Y = \beta_2 Z + u$ , but  $Z$  is measured with measurement error  $w$ , the observed variable being  $X = Z + w$ . It may be assumed that  $w$  has zero mean and constant variance, and that it is distributed independently of  $Z$ . Derive an expression for the limiting value of the usual estimator of the slope coefficient,  $\sum_{i=1}^n X_i Y_i / \sum_{i=1}^n X_i^2$ .

**(2)\*** Exercise 8.8 in the text

Suppose that the true model is  $Y = \beta_2 Z + u$ , but  $Z$  is measured with measurement error  $w$ , the observed variable being  $X = Z + w$ . It may be assumed that  $w$  has zero mean and constant variance, and that it is distributed independently of  $Z$ .  $\bar{Y}/\bar{X}$  is proposed as an estimator of  $\beta_2$ . Derive an expression for its limiting value.

**(3)\*** Exercise 8.14 in the text

It is possible that the *ASVABC* test score is a poor measure of the kind of ability relevant for earnings. Accordingly, perform an OLS regression of the logarithm of hourly earnings on *S*, *EXP*, *ASVABC*, *MALE*, *ETHBLACK*, and *ETHHISP* using your *EAEF* data set and an IV regression using *SM*, *SF*, *SIBLINGS*, and *LIBRARY* as instruments for *ASVABC*. Perform a Durbin–Wu–Hausman test to evaluate whether *ASVABC* appears to be subject to measurement error.

*Stata commands*

The instrumental variables estimation command in Stata is `ivregress 2sls`. It should be followed by the dependent variable, then the explanatory variables not requiring instrumentation, then, in parentheses, the variables requiring instrumentation, followed by an = sign and a list of the instrument(s). Thus the command for an IV regression of *LGEARN* on *S*, *ASVABC*, *MALE*, *ETHBLACK*, and *ETHHISP*, with *SM* instrumenting for *ASVABC*, is

```
ivregress 2sls LGEARN S EXP MALE ETHBLACK ETHHISP (ASVABC = SM SF
SIBLINGS LIBRARY)
```

To perform the Hausman test, continue with the following commands:

```
estimates store name1
reg LGEARN S ASVABC EXP MALE ETHBLACK ETHHISP
estimates store name2
hausman name1 name2, constant
```

'name1' and 'name2' are names to be supplied by you to identify the IV and OLS regressions in question.

For the class in the week beginning January 30:

## PROBLEM SET 12

*Note:* As preparation for Problem Set 14, the class teacher should assign one of the following categories of consumer expenditure, other than *FOOD* or *HOUS*, to each member of the class.

### Expenditure on nondurables

<i>CLOT</i>	Clothing and shoes
<i>FLOW</i>	Flowers, seeds, and potted plants
<i>FOOD</i>	Food purchased for off-premise consumption (this category should not be assigned for practical work because it is used for examples in the text).
<i>GASO</i>	Gasoline and oil
<i>MAGS</i>	Magazines, newspapers, and sheet music
<i>TOB</i>	Tobacco products
<i>TOYS</i>	Nondurable toys and sport supplies

### Expenditure on services

<i>ADM</i>	Admissions to specified spectator amusements
<i>BUSI</i>	Personal business
<i>DENT</i>	Dentists
<i>DOC</i>	Physicians
<i>GAS</i>	Gas
<i>HOUS</i>	Housing (this category should not be assigned for practical work because it is used for examples in the text).
<i>LEGL</i>	Legal services
<i>MASS</i>	Local transportation: mass transit systems
<i>REL</i>	Religious and welfare activities
<i>TELE</i>	Telephone and telegraph

### Expenditure on durables

<i>BOOK</i>	Books and maps
<i>FURN</i>	Furniture
<i>OPHT</i>	Ophthalmic products

The data file Demand.wf1 is described in Appendix B of the text. It contains annual expenditure and price data for the categories of expenditure listed above for the USA for the period 1959–2003. You will be working only with the category assigned to you.

### (1)\* Exercise 9.10 in the text

In principle *ASVABC* might be a positive function of *S*, in which case the educational attainment model should have two equations:

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + u_S$$

$$ASVABC = \alpha_1 + \alpha_2 S + u_A.$$

Using your *EAEF* data set, fit the second equation, first using OLS, second using instrumental variables estimation with *SM* as an instrument. Investigate analytically the likely direction of the bias in the slope coefficient in the OLS regression, demonstrate that in principle IV should yield consistent estimates, and check whether a comparison of the OLS and IV estimates confirms your analysis.

### *Stata commands*

```
reg ASVABC S
ivregress 2sls ASVABC (S = SM)
```

**(3)** Exercise 9.5 in the text

Consider the model

$$Y = \beta_1 + \beta_2 X + u$$

$$X = \alpha_2 Y + v$$

where  $u$  and  $v$  are identically and independently distributed disturbance terms with zero means.

Demonstrate that the OLS estimator of  $\alpha_2$  is, in general, inconsistent. How is your conclusion affected in the special case  $\beta_2 = 0$ ? How is your conclusion affected in the special case  $\alpha_2 \beta_2 = 1$ ? How would your answers to these questions be affected if  $u$  and  $v$  were not distributed independently of each other?

**(3)** Exercise 9.12 in the text

A researcher believes that a model consists of the following relationships:

$$Y = \alpha_1 + \alpha_2 X + u \quad (1)$$

$$X = \beta_1 + \beta_2 Y + v \quad (2)$$

$$Z = \gamma_1 + \gamma_2 Y + \gamma_3 X + \gamma_4 Q + w \quad (3)$$

where  $u$ ,  $v$ , and  $w$  are disturbance terms that are drawn from fixed distributions with zero mean. It may be assumed that they are distributed independently of  $Q$  and of each other. All the parameters may be assumed to be positive and it may be assumed that  $\alpha_2 \beta_2 < 1$ . It may be assumed that the model is correctly specified.

- State which variables in this model are endogenous and which are exogenous.
- The researcher fits (2) using OLS. Evaluate whether the estimate of  $\beta_2$  is likely to be biased. If it is biased, determine the direction of the bias.
- The researcher fits (3) using OLS. Determine whether the parameter estimates are likely to be biased. (You are not expected to evaluate the direction of the bias, if any.)
- The researcher decides to fit (2) using IV, with  $Q_i$  as an instrument for  $Y_i$ . Determine whether he is likely to obtain a consistent estimate of  $\beta_2$ .

For the class in the week beginning February 6:

**PROBLEM SET 13****(1)** Exercise 10.4 in the text

What are the factors influencing going to college? Using your *EAEF* data set, define a binary variable *COLLEGE* equal to 1 if  $S > 12$  and 0 otherwise. Regress *COLLEGE* on *ASVABC*, *SM*, *SF*, and *MALE* (1) using ordinary least squares, and (2) using logit analysis. Calculate the marginal effects in the logit analysis and compare them with those obtained using OLS.

*Stata commands*

```
gen COLLEGE = 0
replace COLLEGE = 1 if S>12
reg COLLEGE ASVABC SM SF MALE
logit COLLEGE ASVABC SM SF MALE
```

**(2)** Exercise 10.6 in the text

Regress the variable *COLLEGE* defined in Exercise 10.1 on *ASVABC*, *MALE*, *SM*, and *SF* using probit analysis. Calculate the marginal effects and compare them with those obtained using OLS and logit analysis.

*Stata command*

```
probit COLLEGE ASVABC SM SF MALE
```

**(3)\*** Exercise 10.16 in the text

The probability density function of a random variable  $X$  is the exponential distribution

$$f(X) = \lambda e^{-\lambda X} \quad \text{for } X > 0$$

with unknown parameter  $\lambda$ . Derive the maximum likelihood estimator of  $\lambda$ , given a sample of  $n$  observations on  $X$ .

**(4)\*** Variation on Exercise 10.17 in the text

A random variable  $X$  can be equal to any of the positive integers in the range 1 to  $n$  with equal probability.  $n$  is unknown. In a sample of 2 observations, the values of  $X$  were 2 and 3. Derive the maximum likelihood estimate of  $n$ . Determine whether it is biased or unbiased.

Demonstrate that  $Z = \text{sum of the two observations}$  is an unbiased estimator. Which estimator should be preferred, the maximum likelihood estimator or  $Z$ ?

*For the class in the week beginning February 13:*

**PROBLEM SET 14****Notes on EViews and the Demand Functions data set**

The Demand Functions data set is a subset of the National Income and Product Accounts published on a regular basis by the US Bureau of the Census. The data set contains annual observations for the period 1959–2003 on aggregate disposable income, 20 categories of consumer expenditure, and price index series for these categories. The income and expenditure variables are all measured in \$ billion at 2000 constant prices. The price index series are all based with 2000=100. The categories of consumer expenditure and their names in the file were listed at the beginning of Problem Set 12. The nominal price index series for each category of expenditure has the name of the category prefixed by *P*: *PFOOD*, *PHOUS*, etc. The data set includes the nominal price index for total personal expenditure, *PTPE*.

*Downloading the data file*

The data file Demand.wf1 may be downloaded to your H:space from the course website. The downloaded file will actually be saved as Demand\_wf1.bin. You must rename it Demand.wf1. You can do this by going to My Computer, H:, File, and Rename. You will download the file only once.

You may find it helpful to simplify the data set by deleting the expenditure and price variables relating to the categories of consumer expenditure other than the one assigned to you. To do this, highlight the variables you wish to discard (you can highlight as many as you like by dragging on the mouse), and click on the Delete tab on the tab bar of the Workfile window. Make sure that you do *not* delete the following:

- the data for your commodity
- the associated price series
- *DPI* (disposable personal income at constant prices)
- *PTPE* (price index for total personal expenditure)

Then save this smaller data set under a new name. In due course you will add further variables to this data set, for example the logarithms of the data for your commodity and the data for *DPI*. When you do this, save the data set again with the same name.

You cannot print your data file directly. If you wish to obtain a print-out of any of the data in the file, you have to enter EViews, click on the Show tab on the Workfile tab bar, type in the name of the variable whose data you wish to list, and click on OK. To print the data, click on the Print tab at the top of the Show window.

### *Printing*

You have four choices for printing data lists or regression results. I recommend the second.

1. You can send material directly to the printer by clicking on the Print tab that appears on the output menu bar.
2. You can block the output with the cursor, copy it (Control-C), and then paste into a Word document. Change the typeface to Courier so that the columns line up properly.
3. You can send it to an ASCII (text) file by clicking on File on the main EViews menu bar and then click on Print setup. You will then be asked to specify a file name. You should include the path to the directory where you want the file located. Note that you have to click on the Print tab every time you want to send the regression results to the output file. I usually forget to do this and then have to do everything a second time. Even a third time. The file will automatically be closed for you when you finish your EViews session. You should choose a different name for each EViews session if you wish to preserve all your regression results.
4. (Also reached via the Print setup menu) You can save the output as frozen output. I have not discovered what this is.

### *Running EViews from the command line*

The instructions in the problem sets assume that you will be using the various windows provided by EViews. However it is also possible to run EViews from the command line, ignoring the windows entirely. Command line instructions are provided for each exercise using EViews.

### *Running an OLS regression*

Click on Quick in the main EViews menu bar, then Estimate Equation. You will now see the equation specification window. Type in  $Y = C + X_1 + X_2 + X_3 + \dots$  where  $Y$  is the dependent variable,  $C$  is the constant, and  $X_1, X_2, X_3, \dots$  are the explanatory variables. You must use this order, and you must use  $C$  for the constant. If, for some reason, you wish to run a regression without a constant, omit  $C$ .

### *Generating new variables*

To generate a new variable from existing ones, click the Genr tab on the Workfile tab bar. In the new window that appears, type in the equation defining the new variable. Examples are given in items (1) and (3) of this problem set. Remember to save the file when you have added new variables, keeping the same name and over-writing the old file.

**(1)** Exercise 11.2 in the text.

Economic theory (and common sense) suggests that one should use real price indices rather than nominal ones in when fitting demand functions, with the real price index defined relative to general inflation as measured by *PTPE*. For example, the real (or relative) price index for food, *PRELFOOD*, is defined as

$$PRELFOOD = 100*(PFOOD/PTPE)$$

Using the Demand Functions data set, construct a relative price index series for your category by dividing its nominal price series by *PTPE*, and multiplying by 100. Plot the series and try to explain why it has changed over the time period.

*EViews commands using EViews windows:*

Click the Genr tab on the Workfile tab bar. In the new window that appears, type in the equation

$$PREL\text{CAT} = 100*(PCAT/PTPE)$$

substituting the name of your commodity for 'CAT'. Do remember to replace 'CAT' with the name of your category of expenditure.

To plot the graph, go to Quick, then to Graph and type in the name of your relative price index; choose line graph.

*EViews commands using the command line:*

```
GENR PREL\text{CAT} = 100*PCAT/PTPE
GRAPH PREL\text{CAT}
```

**(2)** Exercise 11.3 in the text.

Regress your category of expenditure on *DPI* and the relative price index series constructed in Exercise 11.2. Give an economic interpretation of the regression coefficients and perform appropriate statistical tests.

*EViews commands using EViews windows:*

In the regression window type in

```
CAT C DPI PREL\text{CAT}
```

replacing 'CAT' with the name of your expenditure category.

*EViews command using the command line:*

```
LS CAT C DPI PREL\text{CAT}
```

**(3)\*** Exercise 11.4 in the text

Regress the logarithm of expenditure on your category on *LGDP* and the logarithm of the relative price series. Give an economic interpretation of the regression coefficients and perform appropriate statistical tests.

*EViews commands using EViews windows:*

Click the Genr tab on the Workfile window tab bar. In the new window that appears, type in the equations

```
LGCAT = LOG(CAT)
LGDP = LOG(DPI)
```

replacing 'CAT' with the name of your category of expenditure. Note that EViews uses LOG, not LN, when defining natural logarithms. To define *LGPRCAT*, click the GENR tab on the Workfile window tab bar and type in

```
LGPRCAT = LOG(PRELCAT)
```

replacing 'CAT' with the name of your expenditure category. To perform the regression, go to the regression window and type in

```
LGCAT C LGDPI LGPRCAT
```

*EViews commands using the command line:*

```
GENR LGCAT = LOG(CAT)
GENR LGPRCAT = LOG(PRELCAT)
GENR LGDPI = LOG(DPI)
LS LGCAT C LGDPI LGPRCAT
```

**(4)** Exercise 11.9 in the text

Using the Demand Functions data set, regress the logarithm of expenditure on your commodity on *LGDPPI*, the logarithm of your relative price series, and those two variables lagged one period. Repeat the experiment adding *LGDPPI(-2)* and the logarithm of the price series lagged two periods. Compare the regression results, paying attention to the changes in the regression coefficients and their standard errors.

*EViews commands using EViews windows*

In the regression window you should run regressions with the following specifications:

```
LGCAT C LGDPI LGPRCAT
LGCAT C LGDPI LGDPI(-1) LGPRCAT LGPRCAT(-1)
LGCAT C LGDPI LGDPI(-1) LGDPI(-2) LGPRCAT LGPRCAT(-1) LGPRCAT(-2)
```

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT
LS LGCAT C LGDPI LGDPI(-1) LGPRCAT LGPRCAT(-1)
LS LGCAT C LGDPI LGDPI(-1) LGDPI(-2) LGPRCAT LGPRCAT(-1) LGPRCAT(-2)
```

**(5)** Fit a reparameterized specification of the model with two lags, as described in Section 11.3 in the text. Compare the estimated long-run coefficient with counterparts in the three specifications in Item (1). See Problem Set 15 for instructions on generating new variables. Generate new variables  $X1 = LGDPI - LGDPI(-1)$ ,  $X2 = LGDPI - LGDPI(-2)$ ,  $P1 = LGPRCAT - LGPRCAT(-1)$ ,  $P2 = LGPRCAT - LGPRCAT(-2)$ . Then

*EViews command using EViews windows*

In the regression window you should run a regression with the following specification:

```
LGCAT C LGDPI X1 X2 LGPRCAT P1 P2
```

*EViews commands using the command line:*

```
LS LGCAT C LGDPI X1 X2 LGPRCAT P1 P2
```

For the class in the week beginning February 20:

### PROBLEM SET 15

**(1)\*** Exercise 11.11 in the text

Using the Demand Functions data set, regress expenditure on your category on income, relative price, and expenditure on your category lagged one period. Repeat the regression with the variables in logarithmic form. Give an interpretation of the regression coefficients, paying attention to the dynamics implicit in the model.

*EViews commands using EViews windows*

Go to the regression window and type in

```
CAT C DPI PRELCAT CAT(-1)
LGCAT C LGDPI LGPRCAT LGCAT(-1)
```

*EViews commands using the command line:*

```
LS CAT C DPI PRELCAT CAT(-1)
LS LGCAT C LGDPI LGPRCAT LGCAT(-1)
```

**(2)** Exercise 11.15 in the text. Come to class prepared to give an answer. A written answer is not required.

Explain whether the adaptive expectations process is likely to be a plausible representation of dynamics in financial markets.

**(3)\*** Exercise 11.17 in the text.

It has been asserted that the distortions caused by the violation of Assumption C.7 part (2) when fitting

$$Y_t = \beta_2 Y_{t-1} + u_t$$

diminish with the size of  $\beta_2$ . Suppose that  $\beta_2 = 0$  and you fit this relationship, not knowing that  $\beta_2 = 0$ . What can be said about (a) the asymptotic properties of  $b_2$ , (b) its finite-sample properties? What difference would it make if the true and fitted models included an intercept, again supposing  $\beta_2 = 0$ ?

For the class in the week beginning February 27:

### PROBLEM SET 16

**(1)\*** Exercise 12.3 in the text

Perform the Breusch–Godfrey and Durbin–Watson tests for autocorrelation for the logarithmic demand function that you fitted in Exercise 11.4. Is there evidence of autocorrelation? If so, what are the implications for the statistical tests you performed?

*EViews commands using EViews windows:*

Go to the regression window and type in

```
LGCAT C LGDPI LGPRCAT
```

To save the residuals as Z, click the Genr tab on the Workfile window tab bar and type in

```
Z = resid
```

Then go to the regression window and type in

```
Z C LGDPI LGPRCAT Z(-1)
```

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT
GENR Z = resid
LS Z C LGDPI LGPRCAT Z(-1)
```

**(2)** Exercise 12.4 in the text

If your regression application allows you to graph or print the residuals from a regression, do this in the case of your logarithmic demand function. Does an inspection of the residuals corroborate the presence (or absence) of autocorrelation found in Exercise 12.3?

*EViews commands using EViews windows*

Go to the regression window and type in

```
LGCAT C LGDPI LGPRCAT
```

Click on Quick and then on Graph. In the series list window, type in `resid`, click on OK and choose line graph. `resid` always contains the residuals from the most recent regression.

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT
GRAPH RESID
```

**(3)** Exercise 12.5 in the text

Perform a logarithmic regression of expenditure on your commodity on income and relative price, first using OLS and then using the option for AR(1) regression. Compare the coefficients and standard errors of the two regressions and comment.

*EViews commands using EViews windows*

In the regression window you should run regressions with the following specifications:

```
LGCAT C LGDPI LGPRCAT
LGCAT C LGDPI LGPRCAT AR(1)
```

In the regression results, the coefficient printed in the line corresponding to AR(1) is an estimate of  $\rho$ .

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT
LS LGCAT C LGDPI LGPRCAT AR(1)
```

*For the class in the week beginning March 5:*

**PROBLEM SET 17**

**(1)\*** Exercise 12.8 in the text

Perform a logarithmic regression of expenditure on your category of consumer expenditure on income and price using an AR(1) estimation technique. Perform a second regression with the same variables but adding the lagged variables as regressors and using OLS. With Breusch–Godfrey and Durbin  $h$  tests, check that the second specification is not subject to autocorrelation.

Explain why the first regression is a restricted version of the second, stating the restrictions, and check whether the restrictions appear to be satisfied by the estimates of the coefficients of the second regression. Perform a common factor test. If the AR(1) model is rejected, and there are terms with insignificant coefficients in the second regression, investigate the consequences of dropping them.

*EViews commands using EViews windows*

In the regression window you should run regressions with the following specifications:

```
LGCAT C LGDPI LGPRCAT AR(1)
LGCAT C LGCAT(-1) LGDPI LGDPI(-1) LGPRCAT LGPRCAT(-1)
LGCAT C LGCAT(-1) LGDPI LGPRCAT
```

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT AR(1)
LS LGCAT C LGCAT(-1) LGDPI LGDPI(-1) LGPRCAT LGPRCAT(-1)
LS LGCAT C LGCAT(-1) LGDPI LGPRCAT
```

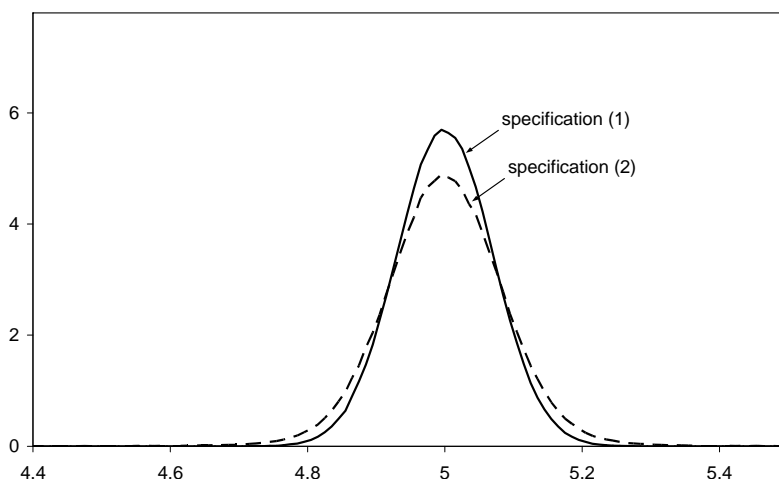
Note that the specification for the third regression is appropriate only for the case where the lagged income and price variables jointly have insignificant explanatory power. This should be tested with an *F* test.

**(2)\*** Exercise 12.11 in the text

Consider the model

$$Y_t = \beta_1 + \beta_2 t + u_t \tag{1}$$

where  $t = 1, \dots, T$  is a time trend, and the disturbance term  $u_t$  is subject to an AR(1) process  $u_t = \rho u_{t-1} + \varepsilon_t$  where  $\varepsilon_t$  is IID with zero mean and finite variance and  $|\rho| < 1$ .



Derive a representation of the time series process for  $Y_t$  that is free from autocorrelation. This will be called specification (2). Explain whether specification (2) may be fitted using OLS. Explain the potential advantages and disadvantages of fitting specification (2) instead of specification (1). Assume that fitting specification (2) using OLS yields consistent estimates of the parameters. Show how one might obtain an estimate of  $\beta_2$  from the regression results and demonstrate that it is consistent.

The figure compares the distributions of the estimates of  $\beta_2$  obtained using specification (1) and those using the method determined via specification (2), using a simulation with the true values chosen as follows:  $\beta_1 = 10$ ,  $\beta_2 = 5$ , and  $\rho = 0.9$ . Comment on the relationship between the distributions in the light of your answers to earlier parts of this exercise.

*For the class in the week beginning March 12:*

### PROBLEM SET 18

**(1)\*** Exercise 13.3 in the text

Consider the model

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 t + u_t$$

and suppose that  $X_t$  has a strong time trend. As a consequence, the estimation of  $\beta_2$  will be subject to multicollinearity and the standard error relatively large. Suppose that one detrends the data for  $Y$  and  $X$  and instead regresses the model

$$e_{Yt} = \beta_2 e_{Xt} + u_t$$

where  $e_{Yt}$  and  $e_{Xt}$  are the residuals from the regressions of  $Y$  on  $t$  and  $X$  on  $t$ . This is a simple regression and so the problem of multicollinearity does not arise. Nevertheless, by virtue of the Frisch–Waugh–Lovell theorem, the estimate of  $\beta_2$  and its standard error will be exactly the same. Explain intuitively why the standard error has not decreased, despite the elimination of the problem of multicollinearity.

**(2)\*** Repeat Granger and Newbold's experiment, with 100 observations instead of 50. Construct two independent 100-observation random walks and regress one on the other. Does the  $t$  statistic on the slope coefficient appear to be significant using a 5 percent test? Repeat the experiment several times (at least 5 times; 10 would be better) and note the frequency of Type I errors.

*EViews commands using EViews windows*

Enter EViews and click on the following: File, New, Workfile.

Choose undated or irregular, keep start observation 1, make the end observation 100.

On the workfile menu, click on the Genr tab and in the window type in

$$Y = 0$$

and then click on OK. Click on Smpl and change the sample period to 2 to 100. Click on Genr a second time, type in

$$Y = Y(-1) + nrnd$$

and click on OK. 'nrnd' is a special function generating a random variable with normal distribution, 0 mean and unit variance. You have now generated a random walk. Change the sample back to 1 to 100 and generate another random walk,  $X$ , in the same way, and regress  $Y$  on  $X$ . Repeat the process at least five times (10 would be better).

*EViews commands using the command line:*

*Note:* You do not have to repeat the first instruction.

```
workfile PS18 u 1 100
GENR Y = 0
SMPL 2 100
GENR Y = Y(-1) + nrnd
SMPL 1 100
GENR X = 0
SMPL 2 100
GENR X = X(-1) + nrnd
SMPL 1 100
LS Y C X
```

*For the class in the week beginning April 23*

## PROBLEM SET 19

(1) Exercise 13.8 in the text.

Perform augmented Dickey–Fuller tests for difference stationarity on the logarithms of expenditure on your commodity and the relative price series. Calculate the first differences and test them for difference stationarity.

*EViews commands using EViews windows*

Open the data file, and click on the name of the variable being tested. Click on the View tab on the series menu, and click on unit root test. A window comes up asking for choices. Keep the Augmented Dickey–Fuller Test, keep Level, and keep lagged differences at 1. Among the options for what to include in the test equation, choose trend and intercept. Click on OK. The ADF test statistic will come up, with the critical values. If you do not reject the null hypothesis that the test is nonstationary, repeat it on first differences (go back to the unit root test choice window and change from Level to 1st difference).

*EViews commands using the command line:*

```
UROOT(T,1) LGCAT
GENR DLGCAT = LGCAT-LGCAT(-1)
UROOT(T,1) DLGCAT
UROOT(T,1) LGPRCAT
GENR DPCAT = LGPRCAT-LGPRCAT(-1)
UROOT(T,1) DPCAT
```

(2)\* Exercise 13.12 in the text

Run logarithmic regressions of expenditure on your commodity on disposable personal income and relative price, plot the residuals, and test for cointegration.

*EViews commands using EViews windows*

Go to the regression window and type in

```
LGCAT C LGDPI LGPRCAT
```

Click the Genr tab and in the window type in

```
Z = resid
```

('resid' always contains the residuals from the most recent regression, but for some reason you cannot use it directly in a test.). Test  $Z$  for nonstationarity. (*Note:* Do not include a time trend in the test specification since there is no reason to expect one.)

*EViews commands using the command line:*

```
LS LGCAT C LGDPI LGPRCAT
GENR Z = resid
UROOT(N,1) Z
```

**(3)\*** Exercise 13.14 in the text

Fit an error correction model for your commodity, assuming that a cointegrating relationship has been found in Exercise 13.12.

*EViews commands using EViews windows*

Click the Genr tab and in the window type in

```
DLGCAT=LGCAT-LGCAT(-1)
```

Repeat this twice to define two other differenced variables:

```
DLGDPI=LGDPPI-LGDPI(-1)
DPCAT=LGPRCAT-LGPRCAT(-1)
```

Go to the regression window and type in

```
DLGCAT Z(-1) DLGDPI DPCAT
```

$Z$  is as defined in Exercise 13.11

*EViews commands using the command line:*

```
GENR DLGCAT = LGCAT-LGCAT(-1)
GENR DLGDPI = LGDPPI-LGDPI(-1)
GENR DPCAT = LGPRCAT-LGPRCAT(-1)
LS DLGCAT Z(-1) DLGDPI DPCAT
```