

1. You have data on the natural log of hourly wages, (LWAGE), age, measured in years, (AGE), a dummy variable, (FEMALE), that takes the value 1 if female 0 otherwise, the product of FEMALE and AGE, FEMALE*AGE, years of work in the current job, TENURE, and its square, TENURE2.

You estimate the following regressions:

(1) reg lhw age female femage tenure tenure2

Source	SS	df	MS			
Model	100.000000	5	16.808223	Number of obs = 3005		
Residual	600.000000	2999	.277909507	F(5, 2299) = 60.48		
				Prob > F = 0.0000		
				R-squared = 0.1429		
				Adj R-squared = 0.1443		
				Root MSE = .52717		

lhw	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0042398	.0014985	2.83	0.005	.0013012	.0071784
female	-.0200000	.0500000	-0.40	0.745	-.1860868	.1331442
femage	-.0100000	.0020022	-5.00	0.000	-.0098752	-.0020224
tenure	.0060000	.0006000	10.00	0.000	.0021002	.0033886
tenure2	-.0005000	.0001000	-5.00	0.000	-.0007000	-.0003000
_cons	1.764674	.0596079	29.60	0.000	1.647783	1.881564

(2) reg lhw age female femage

Source	SS	df	MS			
Model	50.0000000	3	18.8903724	Number of obs = 3005		
Residual	650.000000	3001	.289562779	F(3, 2301) = 65.24		
				Prob > F = 0.0000		
				R-squared = 0.0714		
				Adj R-squared = 0.0772		
				Root MSE = .53811		

lhw	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
age	.0084819	.0014144	6.00	0.000	.0057082	.0112556
female	-.049775	.082544	-0.60	0.547	-.2116433	.1120934
femage	-.0056949	.0020203	-2.82	0.005	-.0096568	-.0017331
_cons	1.762608	.0586584	30.05	0.000	1.647579	1.877637

i) After how many years of work in the current job will the (log of) wages be maximised?
(4 marks)

ii) Test the hypothesis that the coefficients on TENURE and TENURE2 are jointly significant in the model
(6 marks)

iii) What would be the effect for the OLS estimates in equation (1) of omitting the variable FEMALE from the regression?
(6 marks)

iv) Outline how you would test the hypothesis that the specification of the variables on the right hand side of (1) were correct
(9 marks)

2. Given the following model estimated with annual data over the period 1950-1999

$$GDP_t = b_0 + b_1 \text{Investment}_t + u_t$$

you suspect the presence of measurement error in the estimate of the annual level of GDP (measured in £billion).

ie $GDP_t^{\text{observed}} = GDP_t^{\text{true}} + e_t$

where e is a (random) error term

i) Outline the consequences of this type of measurement error for OLS estimation

(6 marks)

ii) given your answer to part i) and the following information, calculate the impact of measurement error in this case

$$\begin{array}{ll} \text{Var}(u) = 2 & \text{Var}(e) = 2 \\ \text{Var}(GDP^{\text{true}}) = 0.2 & \text{Var}(GDP^{\text{observed}}) = 0.5 \\ \text{Var}(\text{Investment}^{\text{true}}) = 0.1 & \text{Var}(\text{Investment}^{\text{observed}}) = 2 \\ \text{Cov}(GDP^{\text{true}}, \text{Investment}^{\text{true}}) = 0.3 & \text{Cov}(GDP^{\text{true}}, \text{Investment}^{\text{observed}}) = 0.2 \\ \text{Cov}(e, u) = 0 & E(u) = 0 \quad E(e) = 0 \end{array}$$

(5 marks)

You are now given new information that says that it is the right hand side variable (Investment) that is instead measured with error

ie
$$\begin{array}{l} \text{Investment}^{\text{observed}} = \text{Investment}^{\text{true}} + w \\ GDP_t^{\text{observed}} = GDP_t^{\text{true}} \end{array}$$

where w is a random error

iii) Find the true (unobserved) OLS estimate of the effect of investment on the level of GDP in the absence of measurement error

(4 marks)

d) the actual OLS estimate given this type of measurement error

(4 marks)

e) Why do the results change like this?

(6 marks)

3. Given the following house price and inflation equations

$$\text{House_Prices}_t = a_0 + a_1 \text{Inflation}_t + u_t \quad (1)$$

$$\text{Inflation}_t = b_0 + b_1 \text{House_Prices}_t + b_2 \text{Interest_Rates}_t + b_3 \text{Money_Supply}_t + e_t \quad (2)$$

a) What would happen if you estimated (1) or (2) by OLS and why? (5 marks)

b) Find the order condition for identification of equations (1) and (2) and say which, if any, instruments, you would use in each case (8 marks)

c) What would be the most efficient solution in equation (1) ? (4 marks)

d) Outline the form of the test to use to check on the endogeneity of any one of the right hand side variables in an equation (8 marks)

TURN OVER

- 4.
- i) What do you understand by the term autocorrelation? (3 marks)
- ii) What can cause it? (4 marks)
- iii) What are the consequences for OLS estimation? (3 marks)

iv) Given the following information from a regression of the model

$$\text{Investment}_t = b_0 + b_1 \text{Investment}_{t-1} + b_2 \text{GNP}_t + b_3 \text{Interest_rates}_t + u_t$$

```
. reg invest invest1 GNP interest
```

Source	SS	df	MS	Number of obs = 29		
Model	1324.46636	3	441.488785	F(3, 25) =	48.69	
Residual	226.700416	25	9.06801664	Prob > F =	0.0000	
				R-squared =	0.8539	
				Adj R-squared =	0.8363	
				Root MSE =	3.0113	
invest	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
invest1	.4749651	.1736102	2.74	0.011	.1174081	.8325221
GNP	.4141392	.1509537	2.74	0.011	.1032442	.7250342
interest	-.2460615	.1169471	-2.10	0.046	-.4869186	-.0052043
_cons	5.412312	2.29867	2.35	0.027	.6781121	10.14651

Durbin-Watson Statistic = 1.589473.

Durbin-Watson h-statistic: .3989117 t = 1.405329

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df
1	2.298	1

H0: no serial correlation

test for the presence of 1st order autocorrelation in the residuals (6 marks)

v) Which of the test statistics do you prefer and why? (4 marks)

vi) Outline the Feasible GLS solution to the problem of autocorrelation (5 marks)

END OF TEST