EE 325 STATA Session II Answers

1. Table 7.11 gives data for the manufacturing sector of the Greek economy for the period 1961-1987.

a. See if the Cobb-Douglas production function fits the data given in the table and interpret the results. What general conclusion do you draw?

Source	SS	df	MS
Model	5.37753949	2	2.68876975
Residual	.158356562	24	.00659819
Total	5.53589605	26	.212919079

Number of obs =	27
F(2, 24) =	407.50
Prob > F =	0.0000
R-squared =	0.9714
Adj R-squared =	0.9690
Root MSE =	.08123

lnoutput	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncapital	.1398108	.1653906	0.85	0.406	-.2015386	.4811603
lnlabor	2.328398	.5994894	3.88	0.001	1.091112	3.565683
_cons	-11.93657	3.211061	-3.72	0.001	-18.56388	-5.30927

b. Now consider the following model:

$$Output / labor = A(K / L)^{\beta} e^{u}$$

where the regressand represents labor productivity and the regressor represents the capital labor ratio. What is the economic significance of such a relationship, if any? Estimate the parameters of this model and interpret your results.

Source	SS	df	MS
Model	2.17537967	1	2.17537967
Residual	.232757738	25	.00931031
Total	2.40813741	26	.09262067

Number of obs =	27
F(1, 25) =	233.65
Prob > F =	0.0000
R-squared =	0.9033
Adj R-squared =	0.8995
Root MSE =	.09649

lnproducti~y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnclratio	.6807562	.0445355	15.29	0.000	.5890337	.7724788
_cons	-1.155956	.0742171	-15.58	0.000	-1.308809	-1.003103

The elasticity of output/ labor ratio (labor productivity) with respect to capital/labor ratio is about 0.68, meaning that if the latter increases by 1% labor productivity, on average, goes up by about 0.68%. A key characteristic of developed economies is a relatively high capital/labor ratio.

2. Table 9.1 Public school teachers' salaries (Dollars) by geographical region. Consider and estimate the following model with the dummy variable:

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i$$

Where Y = (average) salary of public school teacher in state i

$$D_{2i} = 1 \; if \; the \; state \; is \; \text{in} \; the \; Northeast \; or \; North \; Central$$
$$= 0 \; otherwise$$
$$D_{3i} = 1 \; if \; the \; state \; is \; \text{in} \; the \; South$$
$$= 0 \; otherwise$$

t = time

D = 1 for observations in 1970-1981

= 0 otherwise

Find the mean salary of public school teacher in the South and the mean salary of public school teacher in the Northeast or North Central.

Source	SS	df	MS
Model	98985177.3	2	49492588.6
Residual	2.1523e+09	48	44839670.6
Total	2.2513e+09	50	45025787.3

Number of obs	=	51
F(2, 48)	=	1.10
Prob > F	=	0.3399
R-squared	=	0.0440
Adj R-squared	=	0.0041
Root MSE	=	6696.2

salary	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d2	1524.099	2363.139	0.64	0.522	-3227.311	6275.509
d3	-1721.027	2467.151	-0.70	0.489	-6681.566	3239.512
_cons	48014.62	1857.204	25.85	0.000	44280.46	51748.77

$$Y_i = 48,014.62 + 1,524.099 D_{2i} - 1,721.027 D_{3i}$$

The mean salary of public school teacher in the South is $(48,014.62 - 1,721.027).
The mean salary of public school teacher in the Northeast or North Central is $(48,014.62 + 1,524.099).

3. The Use of Dummy Variables in Seasonal Analysis

Table 9.4 U.S. Refrigerator Sales (thousands), 1978-1985 (quarterly)

a. Consider and estimate the following model with the dummy variable:

$$Y_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + u_t$$

Where Y = refrigerator sales, thousands

Source	SS	df	MS
Model	59654886.6	4	14913721.7
Residual	806142.375	28	28790.7991
Total	60461029	32	1889407.16

Number of obs	=	32
F(4, 28)	=	518.00
Prob > F	=	0.0000
R-squared	=	0.9867
Adj R-squared	=	0.9848
Root MSE	=	169.68

frig	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d1	1222.125	59.99041	20.37	0.000	1099.24	1345.01
d2	1467.5	59.99041	24.46	0.000	1344.615	1590.385
d3	1569.75	59.99041	26.17	0.000	1446.865	1692.635
d4	1160	59.99041	19.34	0.000	1037.115	1282.885

b. Instead of assigning a dummy for each quarter and suppressing the intercept term to avoid the dummy variable trap, what should we do?

Source	SS	df	MS		
Model	915635.844	3	305211.948	Number of obs =	32
Residual	806142.375	28	28790.7991	F(3, 28) =	10.60
				Prob > F =	0.0001
Total	1721778.22	31	55541.2329	R-squared =	0.5318
				Adj R-squared =	0.4816
				Root MSE =	169.68

frig	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
d2	245.375	84.83926	2.89	0.007	71.58966	419.1603
d3	347.625	84.83926	4.10	0.000	173.8397	521.4103
d4	-62.125	84.83926	-0.73	0.470	-235.9103	111.6603
_cons	1222.125	59.99041	20.37	0.000	1099.24	1345.01

4. Table 8.9 Savings and Personal Disposable income (billions of dollars), United States, 1970-1995.

a. Given the data in the table, estimate the following linear savings function using personal disposable income

Time period 1970-1981: $Y_t = \lambda_1 + \lambda_2 X_t + u_{1t} \quad n_1 = 12$

Time period 1982-1995: $Y_t = \gamma_1 + \gamma_2 X_t + u_{2t} \quad n_2 = 14$

Time period 1970-1995: $Y_t = \alpha_1 + \alpha_2 X_t + u_{3t} \quad n_1 + n_2 = 26$

Time period 1970-1981: $Y_t = \lambda_1 + \lambda_2 X_t + u_{1t} \quad n_1 = 12$

Source	SS	df	MS		
Model	16456.2587	1	16456.2587	Number of obs =	12
Residual	1785.03254	10	178.503254	F(1, 10) =	92.19
				Prob > F =	0.0000
Total	18241.2912	11	1658.2992	R-squared =	0.9021
				Adj R-squared =	0.8924
				Root MSE =	13.361

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.0803319	.0083665	9.60	0.000	.0616901	.0989737
_cons	1.016115	11.63771	0.09	0.932	-24.91432	26.94655

Time period 1982-1995: $Y_t = \gamma_1 + \gamma_2 X_t + u_{2t} \quad n_2 = 14$

Source	SS	df	MS		
Model	2614.39647	1	2614.39647	Number of obs =	14
Residual	10005.2214	12	833.768451	F(1, 12) =	3.14
				Prob > F =	0.1020
Total	12619.6179	13	970.739837	R-squared =	0.2072
				Adj R-squared =	0.1411
				Root MSE =	28.875

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.0148624	.0083932	1.77	0.102	-.0034248	.0331496
_cons	153.4947	32.71227	4.69	0.001	82.22075	224.7686

Time period 1970-1995: $Y_t = \alpha_1 + \alpha_2 X_t + u_{3t} \quad n_1 + n_2 = 26$

Source	SS	df	MS		
Model	76621.7867	1	76621.7867	Number of obs =	26
Residual	23248.3	24	968.679166	F(1, 24) =	79.10
				Prob > F =	0.0000
Total	99870.0867	25	3994.80347	R-squared =	0.7672
				Adj R-squared =	0.7575
				Root MSE =	31.124

savings	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
income	.0376791	.0042366	8.89	0.000	.0289353	.046423
_cons	62.42267	12.76075	4.89	0.000	36.08578	88.75957

b. On the basis of the Chow test that there was a difference in the regression of savings on income between the two periods. Consider and estimate the following model with the dummy variable:

$$Y_t = \alpha_1 + \alpha_2 D_t + \beta_1 X_t + \beta_2 (D_t X_t) + u_t$$

Where Y = Savings

X= Personal disposable income

t = time

D = 1 for observations in 1982-1995

= 0 otherwise

(i) Test the coefficients individually statistically significant at the 5 percent level? From this test, how would you describe the difference in the two regressions (coincident regression, parallel regression, concurrent regression, dissimilar regression)?

```
      Source |       SS       df       MS              Number of obs =      26
-------------+------------------------------           F(  3,    22) =   54.78
       Model |  88079.8327     3  29359.9442           Prob > F      =  0.0000
    Residual |  11790.2539    22  535.920634           R-squared     =  0.8819
-------------+------------------------------           Adj R-squared =  0.8658
       Total |  99870.0867    25  3994.80347           Root MSE      =   23.15

------------------------------------------------------------------------------
     savings |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           d |   152.4786   33.08237     4.61   0.000     83.86992    221.0872
      income |   .0803319   .0144968     5.54   0.000     .0502673    .1103964
          dx |  -.0654694   .0159824    -4.10   0.000     -.098615   -.0323239
       _cons |   1.016115   20.16483     0.05   0.960    -40.80319    42.83542
------------------------------------------------------------------------------
```

$H_0 : \alpha_2 = 0$

$H_1 : \alpha_2 \neq 0$

$P - value = 0.000$

$reject \ H_0$

$H_0 : \beta_2 = 0$

$H_1 : \beta_2 \neq 0$

$P - value = 0.000$

$reject \ H_0$

There is enough evidence suggest that the coefficients individually statistically significant at the 5 percent level.

The two regressions are dissimilar regressions.

(ii) Write down the mean personal savings function for 1970-1981 and the mean personal savings function for 1982-2005.

Mean savings function for 1970-1981

$$E(Y_t \mid D_t = 0, X_t) = 1.1061 + 0.0803 X_t$$

Mean savings function for 1982-1995

$$E(Y_t \mid D_t = 1, X_t) = (1.0161 + 152.4786) + (0.0803 - 0.0655)X_t = 153.4947 + 0.0148X_t$$

5. Table 10.5 Hypothetical Data on Consumption Expenditure in Relation to Income and Wealth. Consider and estimate the following model:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i}$$

Where

$Y_i = Consumption\ \ Expenditure(\$)$
$X_{2i} = Income(\$)$
$X_{3i} = Wealth(\$)$

i. Estimate the parameters of this model using the data

```
      Source |       SS       df       MS              Number of obs =      10
-------------+------------------------------           F(  2,     7) =   92.40
       Model |  8565.55407     2  4282.77704           Prob > F      =  0.0000
    Residual |  324.445926     7   46.349418           R-squared     =  0.9635
-------------+------------------------------           Adj R-squared =  0.9531
       Total |        8890     9  987.777778           Root MSE      =   6.808

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |   .9415373   .8228983     1.14   0.290    -1.004308    2.887383
          x3 |  -.0424345   .0806645    -0.53   0.615    -.2331757    .1483067
       _cons |   24.77473     6.7525     3.67   0.008     8.807609    40.74186
------------------------------------------------------------------------------
```

ii. Do you expect to face the multicollinearity problem? Why?

```
             |      x2       x3
-------------+------------------
          x2 |  1.0000
          x3 |  0.9990   1.0000
```

Yes. You do expect to face the multicollinearity problem.

High R-Squared but all variables are insignificant.

High pair-wise correlations among regressors.

iii. If you do expect to face the multicollinearity problem, how will you go about resolving the problem?

Dropping a variable

Source	SS	df	MS		
				Number of obs =	10
				F(1, 8) =	202.87
Model	8552.72727	1	8552.72727	Prob > F =	0.0000
Residual	337.272727	8	42.1590909	R-squared =	0.9621
				Adj R-squared =	0.9573
Total	8890	9	987.777778	Root MSE =	6.493

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x2	.5090909	.0357428	14.24	0.000	.4266678	.591514
_cons	24.45455	6.413817	3.81	0.005	9.664256	39.24483

Source	SS	df	MS		
				Number of obs =	10
				F(1, 8) =	176.67
Model	8504.87666	1	8504.87666	Prob > F =	0.0000
Residual	385.123344	8	48.1404181	R-squared =	0.9567
				Adj R-squared =	0.9513
Total	8890	9	987.777778	Root MSE =	6.9383

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x3	.0497638	.003744	13.29	0.000	.0411301	.0583974
_cons	24.41104	6.874097	3.55	0.007	8.559349	40.26274