EE 325 STATA Session PART 2 April 5th, 2012

1. 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* in *the Northeast or North Central*

$= 0$ *otherwise*

$D_{3i} = 1$ *if the state is* 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.

2. **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

$D_{1t} = 1$ *if in the First quarter*

$= 0$ *otherwise*

$D_{2t} = 1$ *if in the Second quarter*

$= 0$ *otherwise*

$D_{3t} = 1$ *if in the Third quarter*

$= 0$ *otherwise*

$D_{4t} = 1$ *if in the Fourth quarter*

$= 0$ *otherwise*

t = time

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

c. Since the models a. and b. do not contain any covariates, will the picture change if we bring in a quantitative regressor in the model?

3. 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: $Y_t = \alpha_1 + \alpha_2 X_t + u_{3t} \quad n_1 + n_2 = 26$

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)?

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

4. 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

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

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