Applied Econometrics Name:

Midterm exam

You can use your notes, the presentations, and the text. You should not work with other people.

Please send your work to me: [derrell@fordham.edu](mailto:derrell@fordham.edu)

The midterm is due Thursday, October 22.

1. The table below gives the estimated coefficients and their p-values for the following regression. The p-values are for the two-sided test of statistical significance. D is a dummy variable.

|  |  |  |
| --- | --- | --- |
|  | Estimated coefficient | P-value |
| X1 | -0.52 | 0.145 |
| X2 | 1.23 | 0.023 |
| D | 0.44 | 0.078 |

a) What is the estimated value of y when x1 = 5, x2 = 10, and D = 1? (Beta0 = 2)

b) Interpret the estimated value of beta1. Is beta1 statistically significant at the 10% level?

*A one unit increase in X1 leads to a 0.52-unit decrease in y. The p-value is above 0.10, so the estimated beta1 is not statistically significant at the 10% level.*

c) Suppose we had thought that beta1 should be negative. Is beta1 significantly negative at the 10% level?

*The p-value for a one-sided test to see if the coefficient is significantly negative is 0.145/2 = 0.0725. This is less than 0.10. The estimated beta is significantly negative at the 10% level.*

d) Interpret the estimated value of beta2. Is beta2 statistically significant at the 5% level?

*A one-unit increase in X2 leads to a 1.23-unit increase in y. The p-value is less than 0.05, so the estimated beta2 is statistically significant at the 5% level.*

e) What is the intercept when D = 0? What is the intercept when D = 1?

*When D = 0, the intercept is 2. What D = 2 the intercept is 2.44.*

f) You are considering the addition of another variable, x4, to the regression. You believe that x4 has a positive effect on y. You think that x4 is positively correlated with x2, but is not correlated with x1 or D. In the above reported regression, which coefficient (or coefficients) might be biased by the omission of x4? What is the direction of the expected bias?

*Only X2 would be biased. X1 and D cannot be biased since they are not correlated with x4.*

*Since the effect of x4 on y and the correlation between x4 and x2 is positive, the bias would be positive. The estimated beta2 is too high.*

2. Estimated regressions

a)

Interpret the slope coefficient in the above estimated regression equation.

*A one-percent increase in x leads to a 0.1822-unit increase in y.*

b)

Interpret the slope coefficient in the above regression equation.

*A one-percent increase in x leads to a 0.147-percent decrease in y.*

c)

In the above regression equation, D is a dummy variable.

What is the estimated slope coefficient when D = 0? *Slope = 1.27*

What is the estimated slope coefficient when D = 1? *Slope = 1.27 – 0.35 = 0.92*

d)

What is the estimated slope coefficient when x = 4?

*Slope = 0.17 + 2(0.05)x = 0.17 + 2(0.05)(4) = 0.57*

3. A scatter plot of y and x shows a u-shaped relationship. Write a regression equation that will fit this relationship.

*This relationship can be estimated with the polynomial form:*

4. You believe x1 and x2 both have positive effects on y. You also believe that the effect of x1 on y grows stronger as x2 increases. Write a regression equation to fit these relationships.

Here we need an interaction term:

5. OLS regression assumes that the mean of the error terms is zero. What must you do to assure that this assumption is met?

*You must estimate the intercept term. The intercept term captures the fixed portion of y that is not explained by the independent variables. If you do not estimate an intercept term (if you force the intercept to be zero), then the fixed portion of y will be found in the error term and the mean of the error term will not be zero.*

6. Consider the following regression equation. Write the null and alternative hypotheses for the test of the overall significance of the regression. Why would you perform this test?

*You might perform this test because of multicollinearity among the independent variables. The regression may have a high R-square but all of the independent variables are individually insignificant. Multicollinearity might be causing this unusual result. By testing the overall significance of the regression we can see if at least one of the slope coefficients is different from zero, though the test does not reveal which ones.*

7. Consider the linear probability model from your last homework assignment. Use the same data. Add the first lag of the dependent variable:

Report the estimated regression equation. Comment on the statistical significance of the slope coefficients. Is this model better than the model without *Recesst-1*? Why or why not?

The p-value of beta1 is <0.0001 in both regressions. The lagged value of spread is statistically significant.

The p-value of beta2 is 0.2905 in the first regression and 0.1126 in the second. The lagged value of Baa is not statistically significant in either regression, though it is very close to the 10% threshold in the second regression. If we test if the coefficient is significantly positive, then the lagged value of Baa is significant at the 10% level in the second regression.

The p-value of beta3 is 0.0470. The lagged value of recess is statistically significant at the 5% level.

Which model is better? Answers will vary. You might notice that the adjusted R-square is higher in the regression including the lagged value of Recess. You might prefer this model because it has a better fit. Or you might prefer the model that includes the lagged value of Recess because its coefficient is statistically significant. Or, you could calculate the percent of correct predictions and choose the model that does better on this measure.