Economics 475: Econometrics
Midterm

Student ID#: Answers

Please answer the following questions to the best of your ability. Remember, this exam is intended to be closed books, notes, programmable calculators, and neighbors. If you have any questions, please raise your hand. Be sure to show your work if you want partial credit. For any hypothesis tests you perform, full points will be given only if the null and alternative hypothesis are stated, the critical value of your test statistic is stated, and your test statistic is correctly computed. All hypothesis test should be done at the 95% level. Points possible are in parenthesis. Feel free to continue your answers on the back of pages. Good Luck!

Roses are red
and OLS is BLUE
except if your model is heteroskedastic
then GLS gives you a clue

1. In honor of Valentine's Day, one might think about the price and quantity of dates (the romantic type, not the fruit). Consider the case of a male undergraduate who wants to go on dates. Being an economics major, this undergraduate considers the price of the date as being the amount he has to pay in order to get a girl to go out with him. Call this P. The quantity of dates available to him in a given month, Q, is a function of P. These variables are explained by the equations:

$$(1) \qquad P_i = \beta_0 + \beta_1 Q_i + \varepsilon_i$$
$$(2) \qquad Q_i = \alpha_0 + \alpha_1 P_i + v_i$$

a. The undergraduate would like to know what happens to his possible number dates as he changes his willingness to pay for a date. Specifically, he wants to know α_1 so he calls many friends and asks them about how many dates they went on last month and how much they paid at each one. Upon gathering this data, he estimates equation (2). What does he find? Be specific—I want to know (if possible) the direction of his estimates and if they are unbiased and/or efficient. (10)

$\uparrow v \rightarrow \uparrow Q \rightarrow \Delta$ in P depends upon $B_1 \rightarrow cov(P,Q) = f(B_1)$

so OLS is biased in a direction dependent upon B_1.

b. At some point, our undergraduate realized that many of the friends he called were at different campuses where the ratio of men to women were different. Let's call this ratio Z_i. Upon realizing that Z_i might be important, he called his friends and asked their campus' level of Z. He wants to employ Z to help him estimate α_1. Does Z help him? Why or why not? (10)

It seems that the Quantity of available dates is a function of Z, or $Q_i = \alpha_0 + \alpha_1 P_i + \alpha_2 Z_i + v_i$

This means that in this system of equations there is 1 exogenous variable (Z_i) and, in this particular equation, there are 2 slope variables. In this case, this equation is underidentified and α_1 is not able to be estimated. The addition of Z does not help.

c. After exploring Z, our intrepid undergraduate also realized that his friends attend campuses that have drastically different spending patterns. For instance, one campus is in downtown Romeo (CO) where it is very inexpensive to go out for a movie and a date. Another campus is in downtown Loveland (CO) where it is very expensive to do the same thing. After another round of phone calling, our undergraduate creates the variable W which measures the cost of a movie and dinner for each observation. Does W help him understand α_1? Why or why not? (10)

In this case W directly influences P so

$P_i = \beta_0 + \beta_1 Q_i + \beta_2 W_i + \varepsilon_i$. Ignoring the Z from part b, this means our system has 1 exogenous variable and since Equation 2 has 1 slope variable, it is exactly identified. The addition of W helps figure out α_1!

2. In class we stated that the correct equation to use when estimating the variance of the slope estimator in a regression with a single independent variable is $var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(X_i-\bar{X})^2}$. Remembering that $\hat{\beta}_1 = \sum c_i(\beta_0 + \beta_1 X_i + \varepsilon_i)$ and $c_i = \frac{X_i-\bar{X}}{\sum(X_i-\bar{X})^2}$, prove $var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(X_i-\bar{X})^2}$. What classical assumptions are needed in order to show this? (20)

$$Var(\hat{B}_1) = E\left[\left(\sum c_i(\beta_0+\beta_1 X_i+\varepsilon_i) - \beta_1\right)^2\right]$$

$$= E\left[\left(\beta_0 \sum c_i + \beta_1 \sum c_i X_i + \sum c_i \varepsilon_i - \beta_1\right)^2\right]$$

$$\sum c_i = 0 \quad , \quad \sum c_i X_i = 1$$

$$= E\left[\left(\beta_1 + \sum c_i \varepsilon_i - \beta_1\right)^2\right]$$

$$= E\left[\left(\sum c_i \varepsilon_i\right)^2\right]$$

$$= E\Big[c_1\varepsilon_1^2 + c_2\varepsilon_2^2 + \cdots + c_n\varepsilon_n^2 + c_1c_2\varepsilon_1\varepsilon_2 + c_1c_3\varepsilon_1\varepsilon_3 + \cdots + c_{n-1}c_n\varepsilon_{n-1}\varepsilon_n\Big]$$

If $E[\varepsilon_1^2] = E[\varepsilon_2^2] = E[\varepsilon_n^2] = \sigma^2$ (homoskedasticity)

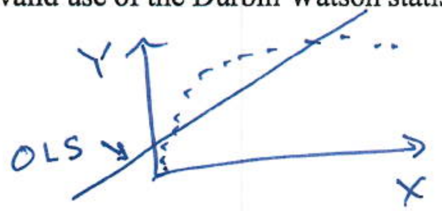
and $E[\varepsilon_i\varepsilon_j] = 0 \quad \forall i \neq j$ (no Autocorrelation)

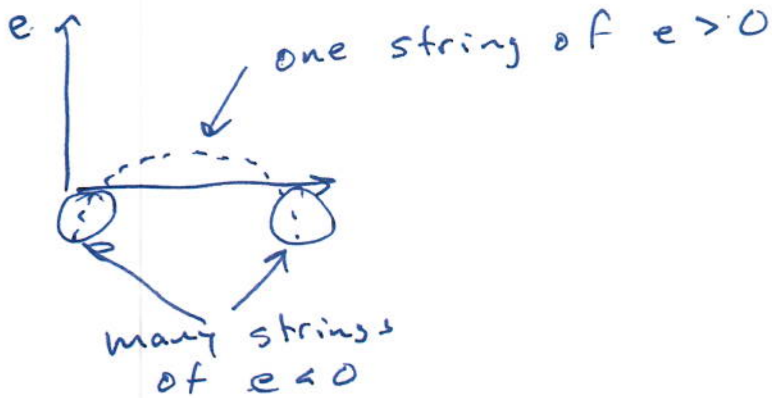
Then $= E[c_1\sigma^2 + c_2\sigma^2 + \cdots + c_n\sigma^2]$

$$= \sigma^2 \sum c_i^2 = \sigma^2 \frac{\left(\sum(x_i-\bar{x})^2\right)}{\left(\sum(x_i-\bar{x})^2\right)^2} = \frac{\sigma^2}{\sum(x_i-\bar{x})^2}$$

3. After taking Dr. Henson's Introduction to Econometrics course, your friend suggests using the Durbin Watson test to test if the unobservable population regression function is nonlinear. Is this a valid use of the Durbin Watson statistic? Explain. (10)

Yes.





4. In a study for Western, Dr. Carl Simpson, a former professor in Sociology, made the claim, "The analytic strategy used in this analysis is Ordinary Least Squares Regression. The dependent variable, salary, is continuous and essentially normally distributed, making regression an efficient and unbiased tool for estimating the unique effects of each factor tested."

Explain why you agree or disagree with this statement. (10)

Disagree. Classical assumptions have nothing to do with the dependent variable.

5. In the third century, Emperor Claudius II ordered that Roman soldiers could not marry. It appears that Claudius believed that unmarried soldiers fought better than married ones.[1] Of course, soldiers wanted to continue in their ways and found a priest, Valentine by name, to marry them irrespective of Claudius's commands. However, an intrepid soldier with a mathematical bent by the name of Cupidius Regressionus decided to test Claudius' theory and measured the military aptitude of married and unmarried soldiers. He then, with a lot of computation, estimated an OLS model of the form:

$$\text{Aptitude}_i = \beta_0 + \beta_1 \text{ Married}_i + \varepsilon_i$$

where Married is a binary variable equal to 1 if the soldier was married and zero otherwise. Cupidius' OLS estimate of β_1 turned out to be statistically significant and positive. Cupidius reported this finding to Claudius, but Claudius, being a knowledgeable emperor, offered a number of criticisms of Cupidius' work. He also ordered Valentine's death.[2] What criticisms could Claudius have offered about this regression? Where possible, provide some solutions that Cupidius could have followed to address these criticisms. (15)

1. Married soldiers have other military skills that unmarried soldiers do not. Since these skills aren't on the right hand side of the regression, they are omitted and $Cov(\varepsilon, X) \neq 0$.

2. Women are attracted to competent men and thus competent men are more likely to be married. This is an example of a system of simultaneous equations and, once again, $cov(\varepsilon, X) \neq 0$.

[1] Interestingly, in 1993 the Commandant of the United States Marine Corps issued almost exactly the same order and tried to prohibit marriage among enlisted Marines.

[2] Also of interest is that Valentine's influence on soldiers is not the only military area connected with him. The British constructed the Valentine Tank during World War II which accounted for about one-quarter of their total tank production. I have to think it felt a little ironic to be a German soldier killed by a Valentine.