Economics 20 Prof. Patricia M. Anderson

Problem Set 3

(due at the beginning of class October **25**, 2002 – note **date change** from syllabus!)

Part I**.** For each of the following questions, choose the one best answer. Briefly explain your reasoning.

1. An economist estimates the following model: $ln(price) = \beta_0 + \beta_1 ln(assess) + \beta_2 ln(lotsize) + \beta_3 ln(sqrft) + \beta_4 bdrms + u$ and plans to use it to predict house prices in his town. He gets predicted ln(price), and then exponentiates it to get a predicted price in dollars. This predicted dollar price:

a) is definitely an underestimate, on average
b) is definitely an overestimate, on average
c) is definitely right, on average
d) impossible to say

2. An economist estimates the following model: $ln(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$ on a sample of 550 workers. He obtains the residuals, *uhat*, and regresses them on *educ*, *exper*, *tenure*, *exper* squared and *tenure* squared. He is carrying out

a) a Breusch-Pagan test for heteroskedasticity
b) a White test for heteroskedasticity
c) an LM test of the null hypothesis that squared terms for *exper* and *tenure* are necessary
d) an LM test of the null hypothesis that squared terms for *exper* and *tenure* are not necessary

3. Suppose an economist wants to estimate the following simple regression model: $y_i = \beta_0 + \beta_1 x_i + u_i$, but decides instead to estimate the following, transformed model: $\frac{y_i}{\sqrt{h_i}} = \beta_0 \frac{1}{\sqrt{h_i}} + \beta_1 \frac{x_i}{\sqrt{h_i}} + \frac{u_i}{\sqrt{h_i}}$.

What must the econometrician think $Var(u_i|x_i)$ is?

a) σ^2/h_i
b) $\sigma^2 h_i$
c) $\sigma^2/\sqrt{h_i}$
d) $\sigma^2 h_i^2$

4. As sample size, n, goes to infinity, then

a) if E(u)=0 and Cov(u,x)=0, the sampling distribution of the OLS estimator converges on the true population parameter
b) the t-statistic becomes normally distributed with mean, 0, and variance 1.
c) the t-statistic becomes t distributed, with n-k-1 degrees of freedom.
d) all of the above

5. Wages tend to be lower in the south than other regions, as seen by the following estimated model: predicted ln(wage) = 1.56 + .05north central + .15west + .12 northeast. If I instead include south and drop north central, then the coefficient on south will be -.05 and

a) all other coefficients stay the same
a) all other coefficients drop by .05
a) all other coefficients increase by .05
a) the intercept increases by .05 and the other dummy coefficients decrease by .05

Part II. Stata Problems. See "Important Things to Know about Stata" on our website for more about installing and using Stata. This PS uses the Stata commands **desc, sum, gen, reg (**and **reg** with the **robust** option**), predict, display** and **Ftail**. All data sets are available through our web page.

1. This problem uses ceosal1.dta, which is a sample CEOs.
a) Look at the variables in the data set. Note that return on stock is measured such that a 25% return is recorded as 25. Create a new variable that records that same variable as .25 instead.

b) Estimate a model of CEO salary as a function of sales and return on stock (in the recoded form). Choose a functional form such that you will directly estimate the elasticity of salary with respect to sales, and that allows the effect of return on stock to possibly be increasing at a decreasing rate. What is the estimated elasticity of salary with respect to sales? Does the effect of return on stock increase at a decreasing rate? What is the turning point? Explain your answer.

c) In the wake of the Enron scandal, you are concerned that the entire CEO salary generating function is different for utility companies versus non-utility companies. Write down the implied (but not necessarily estimated) unrestricted model and the null hypothesis regarding it that this concern would imply testing.

d) Use a Chow test to carry out a test of this null hypothesis. What is the F-statistic? What are the degrees of freedom? What is the p-value associated with it? Do you conclude that utility companies are different or not? Explain. (Feel free to check your calculations by having Stata perform a standard F-test, although it is not necessary).

2. This problem uses nbasal.dta, which is a sample of NBA players.
a) Rick makes a bet with Shaquille that NBA forwards average **more** points per game than do NBA centers. To settle the bet, an economist plans to estimate a model of points per game as a function of position, years of experience (both in the NBA and in college), and talent as measured by draft position. Write down the appropriate null and alternative hypotheses to settle this bet and estimate the model. Assuming a 5% significance level, who would win the bet, based on these estimates? Explain.

b) Before reporting back, you realize that you should consider the possibility of heteroskedasticity. Carry out a Breusch-Pagan test on the estimated model. Can you reject the null of homoskedasticity?

c) To be safe, you decide to re-estimate the model with standard errors that are robust to arbitrary forms of heteroskedasticity. Again assuming a 5% significance level, who would win the bet based on these robust estimates?