Second Assignment, Darren Grant, ECON 3318, Summer 2006. Work the problems, in the format described on the first assignment, on a floppy disk or CD clearly labeled with your name and turn in to me by the due date on the schedule: July 24 at 8:00 am. Assignments turned in after 8:10 am are considered late and will be penalized. Your work should be independent except for minor computer questions, which you may ask of others, and must be in the proper format.

<u>Written Problems</u>. The same rules apply as for assignment 1.

Just for practice: Ch. 3 #2, 8, 9, 10; Ch. 4 #2, 3; Ch. 5 #2, 4, 5, 6, 9, 11.

To turn in: Ch. 3 #3; Ch. 4 #7; Ch. 5 #13.

<u>Computer Problems</u>. The same formatting requirements apply as for assignment 1.

1. (BillTed) Working in the year 2914, archeologists Bill and Ted discover two amazing items from the year 1997. One is called *Centigrade* and the other is called *Fahrenheit*. "What is this crazy equipment?" asks Bill. "I do not know," replied Ted. "Let us try immersing them into pools of water and see what they say." "Excellent!" responds Bill. The two archeologists immerse both of these puzzling pieces of equipment into several pools of water–both in one pool, then both in another pool, over and over again a total of ten times, and record the numbers they get. They then decide to run a regression to figure out how the numbers on *Centigrade* are related to the numbers on *Fahrenheit*.

If the two thermometers were perfectly accurate, they would discover the true, exact linear formula that relates Centigrade and Fahrenheit temperatures, which is F = ___________. Unfortunately, the Fahrenheit thermometer is damaged from being in the ground so long. On average (over many readings) it gives the correct temperature, but each reading is taken with error–on average, an error of ten degrees. Furthermore, the errors are normally distributed (how convenient).

A) On a spreadsheet, choose ten different temperatures between 0 (ice) and 100 (steam). Calculate the corresponding (actual) Fahrenheit temperatures, then regress Fahrenheit (the dependent variable) on Centigrade. Interpret the coefficients and R^2. Pretty good, huh?

B) Continue with the same Centigrade temperatures, but this time add in normally distributed error with a standard deviation of 10 to your Fahrenheit temperature. This can be done using the Random Number Generator in the Analysis Toolpack. Re-estimate your regression and record the coefficient. Is it the same as the coefficient in part A? If not, why not? Is your coefficient estimate biased?

C) Re-do part B ten times with different random numbers. (The way you do this is by changing the seed in the Random Number Generator.) You do not have to record each set of random numbers, but please do record each coefficient estimate, standard error, and t-statistic. Are some coefficient estimates closer to the "true value" than others?

2. (House) This assignment uses cross-section data (Ch. 1) on house prices and characteristics to estimate the "implicit prices" of housing characteristics.

A) It turns out (go figure) that you can obtain appraised values, and house characteristics, for houses of your choosing in Tarrant County. The Tarrant Appraisal District, www.tad.org, allows you to look up the estimated market value (according to the property appraiser) of any property in Tarrant County.

Identify a neighborhood in Tarrant County, close to your house if possible, and identify at least one house by its street address. Then look up market value, age, house size, and the presence of a pool for each of 25 houses in that one neighborhood on the web, and record the data in a spreadsheet. (Sticking to one neighborhood removes bias from unobserved "neighborhood effects," since these aren't measured in the TAD data.) One easy way to look up all the houses in a neighborhood is to look up one house, then click on "georeference."

B) Regress price on age, house size, and a pool dummy. (If none of your houses–or all of them–have a pool, what do you have to do?) Interpret the coefficients. For each, test the null hypothesis that it is equal to zero at the 5% level. Are the coefficient estimates realistic?

3. (BLS) Go to the Bureau of Labor Statistics web site (www.bls.gov) and pull down labor ("hours of all persons"), capital (capital services), output (sectoral output), and multifactor productivity from their productivity data. (You'll have to go to the multi-factor productivity section–each of these is given as an index.) Your data should be annual, span 1950-2001 (the latest data goes to only 2001), and be for the manufacturing sector.

A) Arrange them into a spreadsheet with columns in this order: year, output, labor, capital, and productivity. Regress output on labor, capital, and productivity. Interpret the coefficients. Are they consistent with theory? For each, test the null hypothesis that it is negative at the 5% level.

B) Calculate the annual percentage changes in all variables. Put them in a new sheet in the same order as before: year, change in output, change in labor, change in capital, and change in productivity. Regress the change in output on the change in labor, change in capital, and change in productivity. Interpret the coefficients. Are they the same as those in part A? For each coefficient, test the null hypothesis that it is negative at the 5% level. What about R^2–is it the same or smaller than in part A? Guess which regression is preferred by most economists?