PROBLEM SET 1

PROBLEMS DUE FRIDAY 1/28

Do problems 1, 3, 6, 7, and 1, 2, 3, 4 from Greene's chapters 3 and 4, respectively, as well as the following problem:

Problem 1:

When dealing with time series data it is often observed that economic variables exhibit *time trends,* i.e. a tendency to grow (positive trend) or decline (negative trend) over time. For example, in the table below, the variable X, the number of deaths of children under age 1 (in thousands) exhibits a negative time trend, while the variable Y, the consumption of beer (in bulk barrels) exhibits a positive time trend.

Year	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
X	60	62	61	55	53	60	63	53	52	48	49	43
Y	23	23	25	25	26	26	29	30	30	32	33	31

The presence of such trends may produce spurious results when trying to estimate the relationship between two or more variables. It is thus common practice to *detrend* the variables first. If a variable appears to grow (or decline) linearly with time, it is reasonable to *fit a linear time trend.* A linear time trend may be fitted to X (or Y) by calculating a LS regression of X (or Y) on a constant and time t. The *detrended values* are then the residuals from that regression.

a. Fitting a trend requires choosing an origin and a unit of measurement for the time variable. For example, if the origin is set at mid-1935 and the unit of measurement is 1 year, then the year 1942 corresponds to $t = 7$, and so forth for the other years. If the origin is set at end-1940 (beginning of 1941) and the unit of measurement is 6 months, then 1937 corresponds to $t = -7$. Show that any computed trend values $\hat{X}_t = \hat{a} + \hat{b}t$ are unaffected by the choice of origin and unit of measurement, where $\hat{a}$ and $\hat{b}$ are the LS estimates. Do $\hat{a}$ and $\hat{b}$ change when we change the origin and unit of measurement?

b. Calculate the sample correlation coefficient between X and Y.and the LS slope coefficient of the regression of X on Y and a constant. Is there a positive or negative linear relationship between the two variables? What is the predicted mean of the number of children deaths if the government prohibited beer consumption? By how much does this number increase or decrease if beer consumption increases by one bulk barrel?

c. Even though we may expect in this example some correlation between X and Y (drunk driving may be correlated with children's deaths), the results from (b) above may be exaggerating the strength of the relationship. To see if this conjecture is correct, calculate the correlation coefficient of the *detrended* values of the two variables.

d. Run the least squares regression of the *detrended* values of X on the *detrended* values of Y (without including a constant in the regression). Interpret your results.

e. What happens if we include a constant in the LS regression above?

f. Run the least squares regression of the *raw* (i.e. non-detrended) values of X on the raw values of Y, a constant and time t. How does the slope coefficient of Y compare to that computed in part (d)?