Third Assignment, Darren Grant, ECON 3318, Summer 2006. The formatting and independent work requirements are the same as for the previous assignments. This is due on Thursday, Aug. 3, at 8:00 am; assignments turned in after 8:10 am will be considered late and will be penalized.

Written Problems.

To Turn In: Ch. 5 #11 a, b, d; Ch. 6 #11; Ch. 8 #11. On part b of Ch. 8 #11, multicollinearity does exist–explain why. On part c of Ch. 8 #11, a quadratic in A is a good idea–what coefficient signs should we expect to observe?

Just for practice: Ch. 6 #2, 6, 9, 12; Ch. 7 #4, 6, 7, 9, 12, 13, 14. Ch. 8 #2, 3, 4, 6, 10, 14; Ch. 9 #11; Ch. 10 #2, 9a.

Computer Problems.

1. (Wages) My web site contains a spreadsheet, wages.xls, with information on the real hourly wages and labor market characteristics of 500 individuals in 1983. The characteristics are the following: union membership, years of tenure on the current job, years of schooling, married, female, total years of work experience, and nonwhite race.

A) Regress wages on the nonwhite variable alone. Use a linear specification, and then, separately, use a semi-log specification in which the log of wages is the dependent variable. For each regression, interpret the coefficient, and test the null that it is equal to zero.

Which of the variables in the data are dummy variables? Why?

Why did I use 500 observations, instead of 25 or 30?

B) Add all of the other variables to the regression, using the semi-log specification. Interpret the nonwhite coefficient and test the null that it is equal to zero. The coefficient is closer to zero in this regression. Why? Which coefficient estimate, this one or the one in the previous regression, is a more convincing test of the presence of racial discrimination in the labor market?

C) Extend the specification in part B to include a quadratic in experience. Interpret the coefficients you get. Is the quadratic preferred to the linear specification statistically? Why or why not?

2. (Housets) In an earlier problem we examined cross section housing data–all observations are taken at the same point in time. This problem uses time series housing data–observations are taken from successive years. My website contains a spreadsheet, housets.xls, with six variables: the year, average house price, income (GDP), population, implicit price deflator, and a long term interest rate. (It builds on that described in Chapter 1.) We are interested in assessing the effect of changes in income on housing prices.

A) Regress price on GDP. Interpret the coefficient and assess its statistical significance.

B) Both prices and GDP are in nominal terms. Convert both to real terms using the price deflator and re-run the regression. Interpret the coefficient. This specification is preferred to the previous one. It is more closely linked to economic theory; specifically, to a competitive model of house prices. Explain why.

C) Convert real GDP to per capita terms, and re-run the regression. Interpret the coefficient.

D) Run another regression, again using real per capita GDP, but this time also including the interest rate. Is this variable relevant? Should it be included? Explain why or why not. Interpret the coefficient on the interest rate. What sign do you expect on the coefficient? Construct appropriate null and alternative hypotheses are test them at the 5% level.

3. (Grades) On my web site is a spreadsheet, grades.xls, that contains the grades of 33 students that took Principles of Micro from me at another university, along with their grades in Principles of Macro–which they must take before Micro–and in nine (count 'em) other classes that businesses students are required to take at that university. We are interested in determining if students who do better in Macro do better in Micro as a result, all else equal. The grades are GPA points–4 is an A, 3 a B, etc.

A) Regress the Micro grade on the Macro grade and the grade for each of the other nine classes–a total of ten independent variables. Interpret the coefficient on the Macro grade. Is it statistically significant? Explain the rationale for including these other grades in the regression.

This regression has multicollinearity problems. How can you tell?

B) One way of ameliorating this problem is, instead of including the grade for each of the other nine classes individually, to average them instead. Then the independent variables in the regression would be the Macro grade and the GPA in the other nine classes, just two variables. Conduct this regression. Interpret each coefficient and test the null that it is negative at the 5% level. Give two reasons why this regression is preferred to the previous one.

In this regression, the two coefficients roughly sum to one. Is this reasonable? Why or why not?