EE 325
Practice IV Hints for all answers

1. A variation of the wage-determination equation is as follows:

$$\widehat{W}_t = 1.073 + 5.288V_t - 0.116X_t + 0.054M_t + 0.046M_{t-1}$$
$$(0.797) \quad (0.812) \quad (0.111) \quad (0.022) \quad (0.019)$$

$$R^2 = 0.934 \quad df = 14$$

Where W = wages and salaries per employee
V = unfilled job vacancies in Great Britain as a percentage of the total number of employees in Great Britain
X = GDP per person employed
M = Import prices
M_{t-1} = Import prices in the previous (or lagged) year
(The estimated standard errors are given in the parentheses)

a. Interpret the preceding equation
Ceteris paribus, a 1 percentage point increase in the job vacancy rate lead on average to about 5.29 pounds increase in the wages and salaries per employee; an increase of about 1 pound GDP per person lead on average to about 12 pence decline in wages and salaries per employee; an increase in import prices in the current year and the previous year lead, on average, to an increase in wages and salaries per employee of about 5 pence.

b. Which of the estimated coefficients are individual statistically significant?
State the null and alternative hypothesis for each coefficient first (See multiple regression analysis chapter)
As in the previous exercise, under the zero null hypothesis the estimated t values for the four explanatory variables are, respectively, 6.51, -1.04, 2.45, and 2.42. All but the second of these t values are statistically significant. (See multiple regression analysis chapter on how to get t values) ☺☺

c. Which of the variables may be dropped from the model? Why?
The X variable may be dropped from the model because it has the wrong sign and because its t value is low, assuming of course that there is no specification error. (This is the common questions when you do your empirical paper in your field classes and seminar class after this semester)

d. Test the overall significance of the observed regression.
State the null and alternative hypothesis (See multiple regression analysis chapter)
Use the F test as follows:

$$F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.934/4}{0.66/14} = 49.53$$

This F value is highly significant; for 4 and 14 numerator and denominator degrees of freedom, the 1% level of significance F value is 5.04.
Reject the null hypothesis. There are at least one coefficient not equal to zero.

2. Problem 7 Chapter 6 page 212-213 Wooldridge, J. M. Introductory Econometrics: A Modern Approach. 6th ed. Thompson: South-Western, 2016.

The second equation is clearly preferred, as its adjusted R-squared is notably larger than that in the other two equations. The second equation contains the same number of estimated parameters as the first and one fewer than the third. The second equation is also easier to interpret than the third.

3. Problem 3 Chapter 7 page 252 Wooldridge, J. M. Introductory Econometrics: A Modern Approach. 6th ed. Thompson: South-Western, 2016.

   State the null and alternative hypothesis for each coefficient first (See multiple regression analysis chapter)

(i) The t statistic on hsize2 is over four in absolute value, so there is very strong evidence that it belongs in the equation. We obtain this by finding the turnaround point; this is the value of hsize that maximizes (other things fixed): 19.3/(2 2.19) ≈ 4.41. Because hsize is measured in hundreds, the optimal size of graduating class is about 441.

(ii) This is given by the coefficient on female (since black = 0); nonblack females have SAT scores about 45 points lower than nonblack males. The t statistic is about –10.51, so the difference is very statistically significant. (The very large sample size certainly contributes to the statistical significance.)

(iii) Because female = 0, the coefficient on black implies that a black male has an estimated SAT score almost 170 points less than a comparable nonblack male. The t statistic is over 13 in absolute value, so we easily reject the hypothesis that there is no ceteris paribus difference.

(iv) We plug in black = 1, female = 1 for black females and black = 0 and female = 1 for nonblack females. The difference is therefore –169.81 + 62.31 = -107.50. Because the estimate depends on two coefficients, we cannot construct a t statistic from the information given. The easiest approach is to define dummy variables for three of the four race/gender categories and choose nonblack females as the base group. We can then obtain the t statistic we want as the coefficient on the black female dummy variable.

4. To study the rate of growth of population in Belize over the period 1970-1992, Mukherjee et al. estimated the following models:
   Model I:

$$\ln\widehat{(Pop)}_t = 4.73 + 0.024t$$
$$\text{t} = (781.25)\ (54.71)$$

   Model II:

$$\ln\widehat{(Pop)}_t = 4.77 + 0.015t - 0.075\, D_t + 0.011 D_t t$$
$$\text{t} = (2477.92)\ (34.01)\ \ (-17.03)\ \ (25.54)$$

   where Pop = population in millions, t = trend variable, $D_t = 1$ for observations beginning in 1978 and 0 before 1978, and ln stands for natural logarithm.
   a. In Model I, what is the rate of growth of Belize's population over the sample periods?

b. Are the population growth rates statistically different pre- and post 1978? How do you know? If they are different, what are the growth rates for 19721977 and 1978-1992.

(*a*) 2.4%.

(*b*) Since both the differential intercept and slope coefficients are highly significant, the levels as well the growth rates of population in the two periods are different. The growth rate for the period before 1978 is 1.5% and that after 1978 it is 2.6% (= 1.5% + 1.1%).