Introduction to Econometrics

Midterm Examination

(Make-Up)

Fall 2005

Please answer all of the questions and show your work. Clearly indicate your final answer to each question. If you think a question is ambiguous, clearly state how you interpret it before providing an answer. Be sure to write your name on your answer books!

1. *(24 points)* Say that in a population the conditional expectation of y given x is given by
$$E(y|x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon,$$
where ε has the property that
$$E(\varepsilon|x, x^2) = 0,$$
and
$$E(\varepsilon^2|x, x^2) = \sigma_\varepsilon^2.$$
A researcher has access to a random sample of N observations drawn from this population. If he (mistakenly) estimates the regression function
$$y_i = \delta_0 + \delta_1 x_i + u_i \tag{1}$$
by Ordinary Least Squares (OLS), what will be the property of his estimator $\hat{\delta}$? In particular,

   (a) Show that, in general, $\text{plim}_{N\to\infty} \hat{\delta}_N \neq \delta$.

   (b) Could you give the researcher some advice that would lead him to doubt the validity of the specification he estimated? For example, if after estimating (1) he estimated
   $$y_i = \delta_0 + \delta_1 x_i + \delta_2 x_i^2 + u_i,$$
   should he find that $\hat{\delta}_2 = 0$? Describe how this hypothesis would be formally tested.

2. *(30 points)* Consider estimation of the linear probability model, which is given by
$$d = X\beta + \varepsilon,$$
where d is a "dummy" variable that takes the value 0 or 1.

(a) If $E(\varepsilon|X) = 0$ for all X, show that

$$p(d = 1|X) = X\beta.$$

(b) Show that the expected value of the OLS estimator, $\hat{\beta}$, is equal to β.

(c) Is $\hat{\beta}$ a linear function of the data? If so, can we use the Gauss-Markov Theorem to prove that $\hat{\beta}$ is the Best Linear Unbiased estimator of β? Why or why not?

3. *(24 points)* You estimate a regression function

$$y = X\beta + \varepsilon,$$

using a random sample of observations of size 204 from the population. You can assume that $E(\varepsilon|X) = 0$ and $E(\varepsilon\varepsilon'|X) = \sigma_\varepsilon^2 I_{204}$. You find that the total sum of squares, SST, is equal to 600, while the sum of square errors, SSE, is given by 200. If the column rank of X is 4

(a) Can you obtain an unbiased estimate of σ_ε^2 from the information given? If so, what is it?

(b) What is the square of the multiple correlation coefficient in this case?

4. *(22 points)* You want to estimate the model

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

using feasible generalized least squares. The disturbance term ε is mean independent of the regressor, and the disturbances are independently distributed across population members. Assume that the variable x_i takes only one of three values, x_a, x_b, and x_c. You can assume that you have access to a large sample of size n. Can you obtain a FGLS estimator under these modeling assumptions? If so, describe in detail how you would go about computing it.