UCLA Economics

Spring 2001

Econometrics Field Exam

Please answer four of the six questions. Use a separate blue book for each question.

1. Suppose that conditional on some covariates X, unemployment durations Y have an exponential distribution with mean $exp(\beta_0 + X'\beta_1)$.

   (a) Describe how you would estimate $\beta = (\beta_0, \beta_1')'$ by maximum likelihood. given a random sample of size N.

   (b) What is the variance of the maximum likelihood estimator?

   (c) Describe how you would test the hypothesis that all the elements of β are equal to zero against the alternative that at least some of them differ from zero. The dimension of X is three. Use a likelihood ratio test. Give the critical value.

   (d) Suppose you estimate β by doing ordinary least squares regression of $log(Y)$ on X. How does the variance of the two estimators for β_1 compare?

   Note that if Z has an exponential distribution with mean λ, then the expected value of $log(Z)$ is $c + log(\lambda)$, where c is a constant whose value does not depend on λ.

2. Consider the following model:

$$Y = X\beta + U,$$

   with instruments Z, so that $E[UZ] = 0$, but $E[XU]$ may differ from zero. The dimension of X is smaller than that of Z.

   (a) Describe the two-stage-least-squares estimator for β.

   (b) Describe the optimal gmm estimator for β.

   (c) Describe the empirical likelihood estimator for β.

   (d) Which estimator is more efficient in this case? You may assume that, conditional on Z, U has mean zero and variance σ^2.

   (e) What other differences are there between the three estimators that may affect your choice in a given situation?

3. Consider a regression model with three sets of regressors x, w, z:

$$E(y|x, w, z) = \alpha x + \beta' w + \gamma z$$

Here suppose that x is the scalar variable of interest, and that the vector of variables w is included in the model. The question that arises is whether the scalar z should be included in the model.

(a) How does the inclusion of z affect the estimated standard error for α? Express the change in the estimated standard error of α in terms of (a) a portion due to the dependence between x and z, and (b) a portion due to whatever additional explanatory power accompanies the inclusion of z in the regression. Indicate the direction of the impact each of these components on the estimated standard error.

Relate the second portion (b) above to measures of $\bar{R}^2$ for the short (i.e. excluding z) and long (i.e. including z) regressions. For what relationship between the $\bar{R}^2$ measures of the short and long regressions will you reliably be able to predict the relationship between the standard errors of α in the short and long regressions?

(b) Explain how you can use regression analysis to isolate the components of x and z that jointly determine the long regression estimates of α and γ. From these, construct an R^2 measure of the loss in the explanatory power of x that occurs when z is introduced into the model.

(c) How can you estimate the omitted variable bias in α when z is omitted? Decompose this omitted variable bias estimator into (a) a portion that reflects the dependence between x and z, and (b) a portion that reflects the dependence between y and z.

Express the first component (a) in terms of the R^2 measure from part 3b. From this, express the estimated omitted variables bias in terms of

- a component due to the loss of explanatory power of x when z is added to the model,
- a component due to the dependence between y and z,
- other components. Identify what these other components are and how they affect the bias.

4. Suppose that the Classical Normal Regression Model is applicable to

$$y_i = x_i\beta + \varepsilon_i \qquad i = 1, ..., n$$

where $E(\varepsilon_i^2) \equiv \sigma^2$ is known. We wish to test the set of J restrictions $R\beta = r$. Show that the Wald, Lagrange Multiplier and Likelihood Ratio test statistics are identical. Is this still true when σ^2 is estimated?

5. Derive the probability limits of the first-difference and the OLS estimators of ϕ in the model:

$$y_{it} = \phi y_{it-1} + \alpha_i + \varepsilon_{it} \qquad i = 1, ..., N; \qquad t = 1, 2, 3$$

from a sample of N observations on (y_{i1}, y_{i2}, y_{i3}). Are they consistent? If not, describe a consistent estimator of ϕ. Let the following assumptions hold:

(a) ε_{it} is uncorrelated with all lags of y_{it} and with α_i.

(b) ε_{it} is homoskedastic over time.

(c) $E(\varepsilon_{it}) = E(\alpha_i) = 0$ for all t.

(d) The correlation between y_{it} and α_i is constant over time.

(e) y_{it} is homoskedastic for all t.

Note that (d) implies mean stationarity and with (e) we have full covariance stationarity of y_{it}.

6. Suppose (x_i, y_i) is an i.i.d. sequence, with x_i a scalar and $E(y_i|x_i) = 1 + x_i + \mathbf{1}(x_i > 1/2)$. Assume $x_i \sim U[0,1]$, the uniform distribution on $[0,1]$. Let $z_i = (1, x_i)$, and let $\hat{\beta}$ denote the OLS coefficients in a regression of y_i on z_i. For the following questions state any additional assumptions that are needed.

(a) Let $\beta = \text{plim}\ \hat{\beta}$. What is β? (Give a numerical value.)

(b) Let $\epsilon_i \equiv y_i - z_i'\beta$. Consider $E(\epsilon_i|x_i)$, $E(x_i\epsilon_i)$, and $E(\epsilon_i)$. Which are equal to zero and which are nonzero?

(c) Interpret the results in part (b).