Final Exam, Advanced Econometrics

Jan. 8th, 2016

You have 150 minutes to finish all the six questions.

1. (15 points) Consider the following IV regression model:

$$y_i = \beta x_i + \varepsilon_i,$$

$$E[z_i \varepsilon_i] = c,$$

where β is the unknown scalar parameter, y_i, x_i, and z_i are the dependent variable, a single independent variable which is endogenous, and a single instrumental variable, respectively. c is a nonzero constant which is known.

   (a) Show that β is identified as

$$\beta = \frac{E[z_i y_i] - c}{E[z_i x_i]}.$$

   What do you have to assume about the relationship between the endogenous regressor and instrument for this identification result to be valid?

   (b) Consider the following estimator of β:

$$\hat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n} z_i y_i - c}{\frac{1}{n}\sum_{i=1}^{n} z_i x_i}.$$

   Assuming that the data are i.i.d. and that the assumption imposed in (a) holds, show the consistency of $\hat{\beta}$.

   (c) Establish the asymptotic normality of $\hat{\beta}$.

2. (15 points) Suppose the regression model is

$$y_t = \alpha + u_t$$

where $E(u_t|x_t) = 0$, $cov(u_j, u_i|x_j, x_i) = 0$ for $i \neq j$ but $var(u_t|x_t) = \sigma^2 x_t^2$, $x_t > 0$.

   (a) Given a sample of observations on y_t and x_t, what is the most efficient estimator of α? what is its variance?

(b) What is the OLS estimator of α, and what is its variance?

(c) Prove that the estimator in question (a) is at least efficient as the estimator in question (b).

3. (15 points) For $T = 2$, consider the following model using panel data,

$$y_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it}, \quad t = 1, 2.$$

Let $\hat{\beta}_{FE}$ and $\hat{\beta}_{FD}$ denote the fixed effects and first difference estimators, respectively.

(a) Show that the FE and FD estimates are numerically identical.

(b) Show that the variance matrix estimates from the FE and FD methods are numerically identical.

4. (15 points) These questions are for discrete choice models.

(a) In a logit model, when the outcome variable is binary, derive the log likelihood function, the likelihood equation and the Hessian matrix.

(b) Explain why the logit (or probit) model is better than the Linear Probability Model.

5. (15 points) A log-likelihood function $\log L(\theta_1, \theta_2)$ is a function of two sets of parameters θ_1 and θ_2. Define $\theta_2^*(\theta_1)$ is the solution of

$$\frac{\partial \log L\{\theta_1, \theta_2(\theta_1)\}}{\partial \theta_2} = 0,$$

and define $\log L^*(\theta_1) = \log L(\theta_1, \theta_2^*(\theta_1))$.

(a) Show that the maximum of $\log L(\theta_1, \theta_2)$ with respect to both θ_1 and θ_2 equals the maximum of $\log L^*(\theta_1)$ with respect to θ_1.

(b) Show that $H^* = H_{11} - H_{12}H_{22}^{-1}H_{21}|_{\theta_2=\theta_2^*(\theta_1)}$ where H^* and H are the Hessians of $\log L^*$ and $\log L$ and H is partitioned in the obvious way. (Hint: you may take the equation $\frac{\partial \log L^*(\theta_1)}{\partial \theta_1} = \frac{\partial \log L(\theta_1,\theta_2)}{\partial \theta_1}|_{\theta_2=\theta_2^*(\theta_1)}$ as given)