UCLA Department of Economics
Second-Year Field Examination in Econometrics
September 2007

Instructions:
Solve all three Parts I-III.
Use a separate bluebook for each part.
You have 4 hours to complete the exam.
Calculators and other electronic devices are not allowed.

GOOD LUCK!!!

PART I (Based on Kyriazidou's course)

Problem 1:

(a) Suppose that

$$y_{it} = 1\{x_{it}\beta + \alpha_i - \varepsilon_{it} \geq 0\} \qquad i = 1, ..., N; t = 1, 2$$

where ε_{it} are unobservable variables distributed independently and identically over time conditional on $(x_{i1}, x_{i2}, \alpha_i)$, x_{it} is a $1 \times k$ vector of observable variables, and α_i is an unobservable individual-specific effect. Discuss identification and estimation of β. Assume that cross-sectional sampling is random. Make sure to mention other important assumptions that are used in the identification and consistent estimation of β.

(b) Now suppose that

$$y_{it} = 1\{\beta y_{it-1} + \alpha_i - \varepsilon_{it} \geq 0\} \qquad i = 1, ..., N; t = 1, 2, 3$$

where ε_{it} are distributed independently and identically over time conditional on α_i and α_i is an unobservable individual-specific effect. Assume that y_{i0} is observed for each i although it is not necessarily generated by the same model as the subsequent y_{it}'s. Discuss identification and estimation of β. Assume that cross sectional sampling is random.

Problem 2:

(a) Describe how you would perform Chamberlain's strict exogeneity test in a linear static panel data model of the form

$$y_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it}$$

Make sure to explain the intuition/idea behind the test, to describe the underlying assumptions and to derive its asymptotic distribution. For simplicity you may assume that x_{it} is scalar.

(b) Discuss how you would perform the same test for the static panel data logit model of the form

$$y_{it} = 1\{x_{it}\beta + \alpha_i + \varepsilon_{it} \geq 0\}$$

(HINT: The strict exogeneity concept need to be strengthened from linear projection to conditional mean independence.)

PART II (Based on Winkelmann's course)

Problem 1: Duration analysis

a) Consider a non-negative random variable T (a duration). Assume that the hazard rate is a step function

$$\lambda(t) = \lambda_1 \text{ for } 0 < t < \tau_1$$
$$\lambda(t) = \lambda_2 \text{ for } t \geq \tau_1$$

Derive the survivor function of this model.

b) Find the expected value of T.

c) What is the survivor function if the hazard rate is instead

$$\lambda(t) = \alpha t^{\alpha-1}\lambda_1 \text{ for } 0 < t < \tau_1$$
$$\lambda(t) = \alpha t^{\alpha-1}\lambda_2 \text{ for } t \geq \tau_1 \qquad ?$$

d) Write down the log likelihood function for model in c), for a random sample of n possibly right censored observations, where d_i is an indicator for censoring ($d_i = 1$ if censored, $d_i = 0$ else).

e) Determine the likelihood function of the competing risk model with two independent destinations, where $\lambda_1(t)$ is the hazard rate of exit to destination 1 and $\lambda_2(t)$ is the hazard rate of exit to destination 2.

Problem 2: Count data models

a) Write down the probability function of a Poisson model with hurdle-at-zero.

b) What are the marginal effects at the extensive and intensive margins, respectively?

c) In what kind of empirical situation would you consider using the Poisson hurdle model rather than the simple Poisson model?

d) How can one introduce unobserved heterogeneity into the Poisson hurdle model?

e) An alternative generalization of the Poisson model is the "zero-inflated" Poisson model. How can one test the hurdle Poisson against the zero-inflated Poisson model?

PART III (Based on Guggenberger's course)

Problem 1:

True/Questionable/False? No points are given for just stating true or false, it is the explanation what counts.

1) A stationary AR(1) process $y_t = \rho y_{t-1} + u_t$ (with u_t iid normal) is ergodic.

2) For the Geweke Porter–Hudak estimator $\widehat{d}$ of the long memory parameter d the number of frequencies m used in the pseudo OLS regression has to satisfy $m/n \to 0$ (where n is the sample size) to ensure that the bias of $\widehat{d}$ goes to 0.

3) If in the regression $y_t = \beta x_t + u_t$, y_t and x_t are stationary, then there cannot be a spurious regression problem and the OLS estimator of β is consistent.

Problem 2:

1) In the linear iid IV model

$$\begin{aligned} y_i &= x_i\theta + u_i, \\ x_i &= z_i\pi + v_i, \end{aligned} \qquad (1)$$

derive the asymptotic distribution of the t statistic (that tests $H_0 : \theta = 0$ versus a two sided alternative) under weak instrument asymptotics $\pi = n^{-1/2}c$ for a fixed constant c. Assume that both θ and π are scalars and that the errors are conditionally homoskedastic.

2) Explain the intuition and implementation of Moreira's (Ecta, 2003) conditional likelihood ratio test for the test $H_0 : \theta = 0$ versus a two sided alternative in model (1).

Problem 3:

The goal is to test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$ for given iid data $X_1, ..., X_n$ with $EX_i = \theta$. The test for H_0 is to reject if $T_n = (\widehat{\theta} - \theta_0)/s(\widehat{\theta}) > c$, where c is picked so that the type I error is α ($\widehat{\theta}$ is a root$-n$ consistent estimator of θ and $s(\widehat{\theta})$ is a consistent estimator of the standard deviation of $\widehat{\theta}$). Compare the following two approaches to do so:

1) Using the non-parametric bootstrap, you generate B bootstrap samples, calculate $\widehat{\theta}^*, s(\widehat{\theta}^*)$ for each resample and then calculate $T_n^* := (\widehat{\theta}^* - \theta_0)/s(\widehat{\theta}^*)$, B times. Let $q_n^*(1-\alpha)$ denote the $100(1-\alpha)\%$ quantile of the empirical distribution of T_n^*. You reject H_0 if and only if $T_n > q_n^*(1-\alpha)$.

2) Using subsampling, you calculate $T_{b,i} = (\widehat{\theta}_{b,i} - \theta_0)/s(\widehat{\theta}_{b,i})$, where $\widehat{\theta}_{b,i}$ and $s(\widehat{\theta}_{b,i})$ are based on the same formula as $\widehat{\theta}$ and $s(\widehat{\theta})$ but instead of using all the data, they only use the data $\{X_i, ..., X_{i+b-1}\}$ for $i = 1, ..., n-b+1$. Here b is the blocksize that satisfies $b/n \to 0$. Let $q_{n,b}(1-\alpha)$ denote the $100(1-\alpha)\%$ quantile of the empirical distribution of $T_{b,i}$ for $i = 1, ..., n-b+1$. You reject H_0 if and only if $T_n > q_{n,b}(1-\alpha)$.

Discuss the power properties of the two tests. Are they consistent?

Problem 4:

Explain briefly how the Dickey and Fuller unit root test works.