Studienzentrum Gerzensee Doctoral Program in Economics
Final Exam in Econometrics

February 2003

Bo Honoré and Mark Watson

Write your Identification Number in the space provided below. (Do not give us your name – just your ID Number.)

Identification Number: ____

There are 180 points on this 180-minute exam. The number of possible points for each question is shown in parentheses preceding the question. Please answer all questions on the exam sheet. If you need additional space use the back of the exam sheet. Feel free to use your notes and any textbooks that you may find useful.

(10 points) 1. Suppose that $y_t = \varepsilon_t + 0.3\varepsilon_{t-1}$, where ε_t is i.i.d. with mean zero and variance equal to 4.

(5 points) (a) Derive and plot the sprectrum of y.

(5 points) (b) Let $x_t = 0.5(y_t + y_{t-1})$. Derive and plot the spectrum of x.

(30 points) 2. Suppose that $y_t = x_t\beta + u_t$, where $u_t = \phi u_{t-1} + \varepsilon_t$, $x_t = e_t + \theta e_{t-1}$, where ε_t and e_t and both i..i.d. with mean zero and variance σ_ε^2 and σ_e^2, respectively, and ε_t and e_j are independent for all t and j. Let $\hat{\beta}$ denote the OLS estimator of β based on a sample of size T.

(10 points) (a) Show that $\sqrt{T}(\hat{\beta}-\beta)\xrightarrow{d} N(0,V)$ and derive an expression for V.

(10 points) (b) Suppose that $T = 100$; $\hat{\beta} = 2.1$; $\frac{1}{100}\sum_{t=1}^{100} x_t^2 = 5$; $\frac{1}{99}\sum_{t=2}^{100} x_t x_{t-1} = 2.5$; $\frac{1}{98}\sum_{t=3}^{100} x_t x_{t-2} = 1.0$; $\frac{1}{100}\sum_{t=1}^{100} \hat{u}_t^2 = 4$; $\frac{1}{99}\sum_{t=2}^{100} \hat{u}_t \hat{u}_{t-1} = 3.6$; $\frac{1}{98}\sum_{t=3}^{100} \hat{u}_t \hat{u}_{t-2} = 3.1$; $\frac{1}{99}\sum_{t=2}^{100} x_t \hat{u}_{t-1} = 0.8$; and $\frac{1}{99}\sum_{t=2}^{100} \hat{u}_t x_{t-1} = 0.2$. Construct a 95% confidence interval for β.

(10 points) (c) An alternative to OLS in this problem is GLS. Explain how you construct the GLS estimator. Is GLS preferred to OLS in this situation? Explain.

(20 points) 3. Suppose that $y_t = x_t\beta + u_t$, where $u_t = u_{t-1} + \varepsilon_t$, $u_0 = 0$ and $x_t = \varepsilon_{t+1}$, where ε_t is i.i.d. with mean zero and variance 1. Let $\hat{\beta}$ denote the OLS estimator of β.

(5 points) (a) Prove that $\frac{1}{T}\sum_{t=1}^{T} x_t^2 \xrightarrow{p} 1$

(10 points) (b) Derive the large-sample distribution of $\hat{\beta}$.

(5 points) (c) Is $\hat{\beta}$ consistent? Explain.

(20 points) 4. Suppose that y_{1t} and y_{2t} are scalar random variables with

$$y_{1t} = x_t + \varepsilon_{1t}$$
$$y_{2t} = x_t + \varepsilon_{2t}$$

where x_t, ε_{1t} and ε_{2t} are mutually independent i.i.d sequences of $N(0,1)$ random variables. A researcher has data on y_{1t} and y_{2t} and would like to use these data to estimate the value of x_t. He proposes the estimator $\hat{x}_t = \frac{1}{2} y_{1t} + \frac{1}{2} y_{2t}$.

(5 points) (a) Compute the mean squared error, $E[(\hat{x}_t - x_t)^2]$.

(7 points) (b) A more general form of the estimator is $\tilde{x}_t = \lambda_1 y_{1t} + \lambda_2 y_{2t}$, where λ_1 and λ_2 are two constants. What values of λ_1 and λ_2 yield the estimator with the smallest mean squared error. (If you are short of time, just tell me how to compute them.)

(8 pts) (c) Suppose that instead of being i.i.d x_t follows that AR(1) process $x_t = 0.5x_{t-1} + a_t$ where a_t is i.i.d N(0,1) and is independent of ε_{1t} and ε_{2t}. How would you estimate x_t using the data y_{1i} and y_{2i} $i = 1, \ldots, T$?

(10 points) 5. Suppose that $y_t = h_t\varepsilon_t$ where ε_t is i.i.d (0,1) and $h_t = \alpha_0 + \alpha_1 y_{t-1}^2$

(5 points) (a) Find $E(y_t y_{t-1}^2)$

(5 points) (b) A colleague suggests that y isn't generated by the ARCH(1) model, but rather is generated by the GARCH(1,1) model. How would test this claim using observation on y_t, $t = 1, \ldots, T$?

Problem 6. (10 points)

Suppose that you have a random sample of (y_i, x_i) of size n from

$$y_i = \exp(x_i'\beta) + \varepsilon_i$$

where $E[\varepsilon_i | x_i] = 0$ and $E[\varepsilon_i^2 | x_i] = \exp(x_i'\beta) + 1$. Find the asymptotic distribution of the nonlinear least squares estimator that minimizes

$$\sum_{i=1}^{n} (y_i - \exp(x_i'b))^2$$

Problem 7. (25 points)

Suppose that

$$y = x'\beta + \varepsilon$$

where ε is independent of x and has a $N(0, \sigma^2)$ distribution.

You have two independent samples, each of size n,

- In sample 1, you observe a random sample of (d, x), where $d = 1\{y > 0\}$.
- In sample 2, you observe a random sample of (y, x) *conditional on* $y > 0$

The aim is to estimate (β, σ).

(a) Find the likelihood function.

(b) Describe how you would find the asymptotic distributions of a modified version of Heckman's two step–estimator of (β, σ), where you

1. first estimate $\alpha = \beta/\sigma$ by Probit on the first sample, and
2. secondly estimate (β, σ) by linear regression of

$$y = x\beta + \sigma\lambda(-x\hat{\alpha}) + u$$

using the second sample.

(note that you are not asked to find the asymptotic distribution. Just to explain how you would find it.)

Problem 8. (30 points)

Suppose that you have a random sample $\{X_i\}_{i=1}^n$ from some distribution with density, f. Consider the following two estimators of f at a point x,

$$\hat{f}_1(x) = \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)$$

and

$$\hat{f}_2(x) = \frac{1}{nh_n}\sum_{i=1}^n L\left(\frac{x - X_i}{h_n}\right),$$

where K is the density for a uniform random variable on the interval $(-2, 2)$, and L is the density for a uniform random variable on the interval $(-1, 3)$. h_n is a deterministic sequence that converges to 0 as $n \to \infty$. You may assume that f is continuous and has as many continuous derivatives as you need.

For a given point, x, how do $\hat{f}_1(x)$ and $\hat{f}_2(x)$ compare in terms of bias and variance when n is large.

(answer to problem 3 continued)

Problem 9. (25 points)

Consider observations of (y_{it}, x_{it}, z_{it}) from the linear panel data model

$$y_{i,t} = x_{it}'\beta + z_{it}'\gamma + \alpha_i + \varepsilon_{it}, \quad t = 1, ..., T, \quad i = 1, ..., N$$

where α_i is an unobserved individual-specific effect.

Suppose that x_{it} is strictly exogenous and z_{it} is predetermined in the sense that

$$\begin{aligned} E[\varepsilon_{it} x_{is}] &= 0 && \text{for all } t \text{ and } s \\ E[\varepsilon_{it} z_{is}] &= 0 && \text{for } s \leq t. \end{aligned}$$

No assumption is made on the relationship between α_i and (x_{it}, z_{it}).

(a) What is the minimum T such that β and γ are identified? Explain

(b) Now assume that in addition

$$E\left[\alpha_i z_{is}\right] = 0 \qquad \text{for all } t.$$

What is the minimum T such that β and γ are identified? Explain