Field Exam 2013: Econometrics

August 12, 2013

Please read carefully

You have to:

- Answer **3** of the following 4 questions.

All questions and all subsections of each question have equal weight. No books, notes, tables, or calculating devices are permitted. You have **180** minutes to answer all three questions.

Please make your answers elegant, that is, clear, concise, and, above all, correct. Good luck!

Question 1

Suppose $\left\{(y_t, x_t)' : 1 \leq t \leq T\right\}$ is an observed time series generated by the co-integrated system

$$y_t = \theta_0 x_t + u_t,$$

where

$$\begin{pmatrix} u_t \\ \Delta x_t \end{pmatrix} \sim i.i.d.\ \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

with initial condition $x_0 = 0$.

It can be shown that

$$\begin{pmatrix} T^{-1/2} x_{\lfloor T\cdot \rfloor} \\ T^{-1} \sum_{t=1}^{T} x_t u_t \\ T^{-3/2} \sum_{t=1}^{T} x_t^2 u_t \end{pmatrix} \to_d \begin{pmatrix} B_x(\cdot) \\ \int_0^1 B_x(r)\, dB_y(r) \\ \int_0^1 B_x(r)^2\, dB_y(r) \end{pmatrix},$$

where B_x and B_y are independent Wiener processes.

Let $z_t = (y_t, x_t)'$ and define the function

$$h_T(z_t,\theta) = \begin{pmatrix} x_t/\sqrt{T} \\ x_t^2/T \end{pmatrix} (y_t - \theta x_t).$$

(a) Show that $\Theta_T = \{\theta_0\}$, where $\Theta_T = \left\{\theta : \sum_{t=1}^T E[h(z_t,\theta)] = 0\right\}$.

Let

$$\hat{\theta}_W = \arg\min_\theta g_T(\theta)' W g_T(\theta), \qquad g_T(\theta) = \frac{1}{T}\sum_{t=1}^T h_T(z_t,\theta),$$

where W is a symmetric, positive definite 2×2 matrix.

(b) It can be shown that

$$T\left(\hat{\theta}_W - \theta_0\right) \to_d \int_0^1 B_W(r)\, dB_y(r),$$

where B_W is some functional of B_x and W. Verify this claim and express B_W in terms of B_x and W.

Let $\omega_W^2 = \int_0^1 B_W(r)^2\, dr$.

(c) Find W^*, a value of W for which ω_W^2 is minimal, and express $\omega_{W^*}^2$ in terms of B_x.

(d) Propose a feasible estimator $\hat{\theta}$ satisfying

$$T\left(\hat{\theta} - \theta_0\right) \to_d \int_0^1 B_{W^*}(r)\, dB_y(r).$$

Question 2

1. A scalar dependent variable y_{ij} for n_j individuals in J groups ($i = 1, ..., n_j$ and $j = 1, ..., J$) is assumed to satisfy a nonlinear model

$$y_{ij} = g(x_j, \theta_0) + \varepsilon_{ij}$$

for some group-specific regressors $x_j \equiv x_{ij}$ and unknown p-dimensional parameter with error terms ε_{ij} that are independent across i and j and satisfy a conditional median restriction

$$\Pr\{\varepsilon_{ij} < 0 | x_j\} = 1/2.$$

The error ε_{ij} is assumed to be continuously distributed conditional on x_j, with conditional densities $f(\varepsilon | x_j)$ that are strictly positive everywhere.

For each group j, you are given the sample median $\hat{m}_j$ of y_{ij}, defined as

$$\hat{m}_j = \arg\min_c \sum_{i=1}^{n_j} |y_{ij} - c|, \qquad j = 1, ..., J.$$

You are also given consistent estimators $\hat{\phi}_j$ of the conditional density of ε_{ij} at zero, i.e., consistent estimators of $\phi_j \equiv f(0|x_j)$, which could be used to construct asymptotic variances for the median estimators $\{\hat{m}_j\}$ in the usual way.

a. Under the assumption that $N = \sum_j n_j \to \infty$ with $\lim(n_j/N) \equiv \pi_j > 0$ for all j, and assuming (for the moment) that the group conditional means $\{\gamma_j\}$ are *known*, derive an expression for the asymptotic distribution of the nonlinear least-squares estimator $\tilde{\theta}$ of θ_0 using the group median estimators, i.e., the asymptotic distribution of

$$\tilde{\theta} = \arg\min_{\theta\in\Theta} \sum_{j=1}^{J} (\hat{m}_j - g_j(\theta))^2,$$

where

$$g_j(\theta) \equiv g(x_j, \theta).$$

You need not explicitly show consistency or verify that the relevant remainder terms are negligible, and you should assume that the derivative vectors

$$G_j \equiv \frac{\partial g(x_j, \theta_0)}{\partial \theta}$$

satisfy the needed rank condition(s) to ensure invertibility of the relevant matrices.

b. Propose an alternative estimator $\hat{\theta}$ of θ_0 that efficiently uses the available data $\{\hat{m}_j, x_j, \hat{\phi}_j, n_j\}$, derive its asymptotic distribution, and explain the sense in which it is efficient.

c. Suppose now that the original observations on y_{ij} were censored above at some known, constant value c_0, i.e.,

$$y_{ij} = \min\{g(x_j, \theta_0) + \varepsilon_{ij}, c_0\}.$$

If you deleted the observations on groups where $\hat{m}_j = c_0$ and used the same nonlinear least-squares estimator as in part *a.* above using the remaining groups, would the resulting estimator still be consistent for the true θ_0? Explain. (Assume $g(x_j, \theta_0) \neq c_0$ for all possible x_j.)

d. Finally, suppose the group-specific regressors $\{x_j\}$ are unknown, but you have consistent estimators $\{\hat{x}_j\}$ of them that are conditionally independent of the $\{\hat{m}_j\}$ given the true regressors $\{x_j\}$, with asymptotic distributions of the form

$$\sqrt{n_j}\,(\hat{x}_j - x_j) \xrightarrow{d} N\left(0, \tau_j^2\right).$$

Derive the asymptotic distribution of the "feasible" nonlinear least-squares estimator

$$\tilde{\theta}_F = \arg\min_{\theta \in \Theta} \sum_{j=1}^{J} (\hat{m}_j - \hat{g}_j(\theta))^2,$$

where now

$$g_j(\theta) \equiv g(\hat{x}_j, \theta).$$

Again, assume the estimator is consistent, the relevant remainder terms are asymptotically negligible, and any relevant rank conditions hold for the derivative vectors $\{G_j\}$ and

$$H_j \equiv \frac{\partial g(x_j, \theta_0)}{\partial x}.$$

Question 3

Suppose $\{y_t : 1 \leq t \leq T\}$ is an observed time series generated by the model

$$y_t = \mu + u_t, \qquad u_t = \rho u_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, T,$$

where $u_0 = 0$ and $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$, while $\mu \in \mathbb{R}$ and $\rho \in (-1,1)$ are (possibly) unknown parameters.

(a) Find the log likelihood function $\mathcal{L}(\mu, \rho)$ and, for $r \in (-1,1)$, derive $\hat{\mu}(r) = \arg\max_{\mu} \mathcal{L}(\mu, r)$, the maximum likelihood estimator of μ when ρ is assumed to equal r.

(b) Find the limiting distribution (after appropriate centering and rescaling) of the "oracle" estimator $\hat{\mu}(\rho)$.

(c) Give conditions on $\hat{\rho}$ under which $\hat{\mu}(\hat{\rho})$ asymptotically equivalent to $\hat{\mu}(\rho)$.

(c) Does $\hat{\rho} = 0$ satisfy the condition derived in (c)? If not, determine whether $\hat{\mu}(0)$ is asymptotically equivalent to $\hat{\mu}(\rho)$.

Question 4

Let $X \in \mathbb{R}$ be a random variable with probability measure given by P_0, suppose that for a given $(x, \theta) \mapsto m(x, \theta)$ (which functional form is known) the following restriction over the true parameter θ_0 holds:

$$E_{P_0}[m(X, \theta_0)] > E_{P_0}[m(X, \theta)],$$

for all $\theta_0 \neq \theta = (\beta, \tau) \in \Theta \subseteq \mathbb{R}^d \times \mathbb{R}^k$. Suppose the econometrician observes a sample $(X_i)_{i=1}^n$ of IID copies of X and her/his purpose is to estimate β_0 (think of τ_0 as a nuisance parameter).

Suppose that

1. θ_0 is in the interior of Θ.

2. $\theta \mapsto m(x, \theta)$ is twice continuously differentiable a.s.-X, with uniformly bounded derivatives; $E[\nabla_\beta m(X, \theta_0) \nabla_\beta m(X, \theta_0)']$ exists; and $\nabla_{\beta\beta} E[m(X, \theta_0)]$ is non-singular. [1]

3. $|n^{-1} \sum_{i=1}^n \{\nabla_\beta m(X_i, \theta) - \nabla_\beta m(X_i, \theta_0) - E_{P_0}[\nabla_\beta m(X, \theta)] + E_{P_0}[\nabla_\beta m(X, \theta_0)]\}| = o_{P_0}(n^{-1/2})$.

Please answer the following questions: [2]

1. Suppose that there exists an estimator of τ_0 given by $\hat{\tau}_n = \tau_0 + o_{P_0}(1)$.

   (a) Show that the "two-step estimator" of β_0, $\hat{\beta}_n = \arg\max_\beta n^{-1} \sum_{i=1}^n m(X_i, \beta, \hat{\tau}_n)$, satisfies the following asymptotic linear representation

$$\sqrt{n}(\hat{\beta}_n - \beta_0) = (\nabla_{\beta\beta} E_{P_0}[m(X, \theta_0)])^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \nabla_\beta m(X_i, \theta_0) + \sqrt{n} E_{P_0}[\nabla_\beta m(X, \beta_0, \hat{\tau}_n)] \right) + o_{P_0}(1).$$

2. Derive the asymptotic linear representation of the M-estimator of β_0 when τ_0 is known, i.e., $\tilde{\beta}_n = \arg\max_\beta n^{-1} \sum_{i=1}^n m(X_i, \beta, \tau_0)$.

3. Under what conditions over $\sqrt{n} E_{P_0}[\nabla_\beta m(X, \beta_0, \hat{\tau}_n)]$ the two representations (the ones you found in # 1 and # 2) are equivalent? Derive the conditions formally and provide intuition.

4. Suppose further that

$$\sqrt{n} E_{P_0}[\nabla_\beta m(X, \beta_0, \hat{\tau}_n)] \Rightarrow N(0, \Omega).$$

   Derive under this condition the asymptotic distribution of the "two-step" estimator, $\hat{\beta}_n$. Compare this result with the asymptotic distribution of the estimator when τ_0 is known, i.e., $\tilde{\beta}_n$. In particular, compare the asymptotic variances: can you rank them?

5. Consider now a third estimator. The estimator is given by

$$\theta_n \equiv (\beta_n, \tau_n) = \arg\max_{\beta, \tau} n^{-1} \sum_{i=1}^n m(X_i, \beta, \tau).$$

[1] Henceforth, $\nabla_\beta m(x, \theta)$ denotes the derivative with respect to β at θ; $\nabla_{\beta\beta} m(x, \theta)$ denotes the second derivative with respect to β at θ; etc.

[2] If you think that you need additional regularity conditions in order to answer the following questions, be explicit about it and state them formally and clearly. There will, however, be penalties for adding redundant conditions.

(a) Show that

$$\sqrt{n}(\theta_n - \theta_0) = \left(\nabla E[m(X,\theta_0)]\right)^{-1} n^{-1/2} \sum_{i=1}^{n} \nabla m(X,\theta_0) + o_{P_0}(1)$$

where $\nabla m(x,\theta) \equiv (\nabla_\beta m(x,\theta), \nabla_\tau m(x,\theta))'$ and

$$\nabla E[m(X,\theta_0)] = \left[\begin{array}{cc} \nabla_{\beta\beta} E[m(X,\theta_0)] & \nabla_{\beta\tau} E[m(X,\theta_0)] \\ \nabla_{\tau\beta} E[m(X,\theta_0)] & \nabla_{\tau\tau} E[m(X,\theta_0)] \end{array} \right].$$

Hint: In order to derive this result you will need to impose new regularity conditions. Be explicit about them.

(b) Derive from the previous point the the asymptotic linear representation of β_n. **Hint:** One might need to use the following fact

$$\left[\begin{array}{cc} A & B \\ C & D \end{array} \right]^{-1} = \left[\begin{array}{cc} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{array} \right]$$

(c) Under what conditions over m (namely its derivatives at θ_0) the asymptotic linear representation of β_n coincides (up to negligible terms) with that of $\tilde{\beta}_n$? Interpret your condition.