Field Exam 2017: Econometrics

August 4, 2017

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

$$y_t = \delta t + u_t, \qquad t = 1, \ldots, T,$$

where $u_t \sim i.i.d.\ \mathcal{N}(0,1)$, while δ is the parameter of interest.

As an estimator of δ, consider

$$\hat{\delta} = \frac{\sum_{t=1}^{T} z_t y_t}{\sum_{t=1}^{T} z_t t},$$

where $\{z_t : 1 \leq t \leq T\}$ is some observed time series, which is independent of $\{u_t : 1 \leq t \leq T\}$.

(a) Suppose $z_t = t$. Find the limiting distribution (after appropriate centering and rescaling) of $\hat{\delta}$.

(b) Suppose $z_t = \varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$. Find the limiting distribution of $\hat{\delta}$.

(c) Suppose $z_t = \sum_{s=1}^{t} \varepsilon_s$, where $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$. Find the limiting distribution of $\hat{\delta}$.

(d) Rank the estimators from (a)-(c) in terms of (asymptotic) efficiency.

Question 2

Let Y, X, ε be three scalar continuous random variables and let Z be a binary random variable. Assume these four variables are related by the equation

$$Y = Z \cdot g(X) + \varepsilon,$$

where the unobservable error term ε is assumed to have many moments and satisfy

$$E[\varepsilon|X] = 0,$$

and where the "structural function" $g(\cdot)$ is a very smooth, measurable function on the support of X.

1. Give an expression for $E[Y|X, Z]$, and compare that expression with the structural function g evaluated at X.If you were given only $E[Y|X, Z = 0]$ and $E[Y|X, Z = 1]$, would they suffice to (point) identify $g(X)$? If so, how? If not, what additional information or restrictions would suffice for point identification of $g(x)$? Absent such additional information or restrictions, what can you conclude about the possible values of $g(X)$?

2. Show that, under a suitable restriction on the joint distribution of Z and X, the identified functions $r(x) \equiv E[Y|X = x]$ and $p(x) = E[Z|X = x]$ can be used to identify $g(x)$, and explicity state this restriction on X and Z.

3. Using a random sample of size n, suppose you are given nonparametric (kernel) estimators $\hat{r}(x)$ and $\hat{p}(x)$ of $r(x)$ and $p(x)$, where both are calculated using a symmetric, nonnegative kernel $K(u) = K(-u)$ that integrates to one, and both use a bandwidth of the form $h_n = h_0 n^{-\alpha}$ for some $\alpha \in (1/5, 1)$. Defining an estimator $\hat{g}(x)$ of $g(x)$ based upon the identification result in (b) above, show that it has the same asymptotic distribution as a kernel regression estimator (using the same kernel and bandwidth sequence) of the conditional mean $E[W|X]$ of a particular random variable W which depends upon the observable variables Y, X, Z and the (true) identified functions $r(x)$ and $p(x)$. You need **not** state regularity conditions nor verify the form of the asymptotic distribution, but you should give an explicit expression for W and an argument for the asymptotic equivalence of $\hat{g}(x)$ and the kernel estimator $\hat{E}[W|X = x]$

Question 3

Suppose $\{(y_{it}, x_{it})' : 1 \le i \le n, 1 \le t \le T\}$ is generated by the panel cointegration model

$$y_{it} = \alpha_i + \beta x_{it} + u_{it},$$

where $\{(y_{it}, x_{it})' : 1 \le t \le T\}$ are independent across i and

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

with initial conditions $x_{i0} = 0$, while $\{\alpha_i : 1 \le i \le n\}$ are unknown nuisance parameters and β is the parameter of interest.

(a) Derive an expression for $\hat{\beta}_{ML}$, the maximum likelihood estimator of β.

(b) Suppose n is "large". Holding T fixed, find the limiting distribution (after appropriate centering and rescaling) of $\hat{\beta}_{ML}$ as $n \to \infty$.

(c) Suppose T is "large". Holding n fixed, find the limiting distribution of $\hat{\beta}_{ML}$ as $T \to \infty$.

(d) Suppose both n and T are "large". Does the "large T" version of the result from (b) agree with the "large n" version of the result from (c)?

Question 4

Consider the following IV non-parametric regression model

$$Y = \theta_0(Z) + U, \ where \ E[U|X] = 0$$

for some $\theta_0 \in L^2 \equiv L^2(\mathbb{Z})$, where X is the "instrument" and Z is the "endogenous" variable.

Assume that the conditional expectation operator — which maps $h \in L^2(\mathbb{Z})$ into $E[h(Z)|X = .] = \int h(z)P(dz|X = .) \in L^2(\mathbb{X})$ —, and $x \mapsto r(x) \equiv E[Y|X = x]$ are *known* to you (the applied researcher). Finally, let $||\theta||_w^2 \equiv E\left[(E[\theta(Z)|X])^2\right]$ be the so-called "weak norm".

1. (i) Show that $\theta_0 \in \arg\min_{\theta \in L^2} ||\theta_0 - \theta||_w^2$, but it may not be unique. (ii) What conditions over the conditional expectation operator are needed to ensure point identification? (iii) What does (i) and (ii) tell you about the relationship between the norms $||.||_w$ and $||.||_{L^2}$.

Let, for any $k \in \mathbb{N}$,

$$\theta_k \equiv \arg\min_{\theta \in L^2} ||\theta_0 - \theta||_w^2 + \lambda_k ||\theta||_{L^2}^2$$

where $\lambda_k > 0$. [1] For any $k \in \mathbb{N}$:

2. Show that θ_k exists (in the sense that the "arg min" is non-empty) and that is unique.

3. Show that $||\theta_k - \theta_0||_w^2 + \lambda_k ||\theta_k||_{L^2}^2 = O(\lambda_k)$.

Consider the estimator for any $k \in \mathbb{N}$,

$$\hat{\theta}_k \equiv \arg\min_{\theta \in L^2} n^{-1} \sum_{i=1}^{n} (r(X_i) - E[\theta(Z)|X_i])^2 + \lambda_k ||\theta||_{L^2}^2$$

5. Show that

$$||\hat{\theta}_k - \theta_k||_w^2 + \lambda_k ||\hat{\theta}_k - \theta_k||_{L^2}^2 = O_P(\delta_n)$$

where $(\delta_n)_n$ is such that $\sup_{\theta \in L^2} |n^{-1} \sum_{i=1}^{n} (r(X_i) - E[\theta(Z)|X_i])^2 - E[(r(X) - E[\theta(Z)|X])^2]| = O_P(\delta_n)$. **Hint:** At one point, you may want to use the fact θ_k satisfies a first order condition and that $||\theta_0 - .||_w^2 + \lambda_k || \cdot ||_{L^2}^2$ is quadratic.

6. i. What type of restrictions on $(\lambda_k)_k$ and $(\delta_n)_n$ are needed to ensure consistency of $\hat{\theta}_k$ to θ_0 *under* $||.||_w$? What is the rate of convergence?

   ii. Answer part i., but using $||.||_{L^2}$ instead of $||.||_w$.

[1] $||\theta||_{L^2}^2 = \int |\theta(z)|^2 P(dz)$.