Econometrics Field Exam

Department of Economics, University of California - Berkeley

August 2019

Please choose three questions to answer. Please use a separate blue book for each answered question. Write your name on each blue book. Answer as completely as you are able. Good luck!

[Q1] A dependent variable y_i is generated by a linear regression equation

$$y_i = x_i'\beta_0 + \varepsilon_i,$$

given an observed p-dimensional regression vector x_i, unobserved error term ε_i, and unknown coefficient vector β_0. The errors ε_i are assumed to be independent of x_i with density function $f(\varepsilon)$ that is assumed to be positive everywhere, uniformly bounded, and smooth (i.e., lots of continuous derivatives), and has zero median, i.e.

$$\Pr\{\varepsilon_i \leq 0\} = 1/2.$$

Given a random sample of size N from this model, consider a penalized version of the usual LAD estimator:

$$\begin{aligned}
\hat{\beta} &= \arg\min_{b \in \mathbb{R}'} P_N(\beta) \\
&\equiv \arg\min_{b \in \mathbb{R}'} \left(\left(\frac{1}{N} \sum_{i=1}^{N} |y_i - x_i'b| \right) + \frac{1}{2N}(b - \delta_0)' A_0 (b - \delta_0) \right) \\
&\equiv \arg\min_{b \in \mathbb{R}'} \left(S_N(b) + \frac{1}{2N}(b - \delta_0)' A_0 (b - \delta_0) \right),
\end{aligned}$$

where $S_N(b)$ is the usual LAD criterion function and δ_0 is a known "prior guess" of the unknown β_0 and A_0 is a known, positive-definite weight matrix.

(a) Under what additional conditions (if any) will $\hat{\beta}$ be consistent for β_0?

(b) Assume that the usual "approximate first-order condition"

$$\sqrt{N} \frac{\partial^- P_N(\hat{\beta})}{\partial b} = o_p(1),$$

where $\partial^- P_N(b)/\partial b$ is the subgradient of $P_N(b)$. Also assume the following "approximate mean-value expansion" holds:

$$\sqrt{N} \frac{\partial^- S_N(\hat{\beta})}{\partial b} = \sqrt{N} \frac{\partial^- S_N(\beta_0)}{\partial b} + H_0 \sqrt{N}(\hat{\beta} - \beta_0) + o_p(1),$$

where H_0 is the appropriate "Hessian" matrix for LAD regression, assumed invertible. Under these additional restrictions, derive the form of the asymptotic distribution of $\hat{\beta}$, assuming it is consistent for β_0. You need not check regularity conditions, but please make your expressions as explicit as possible (including the correct expression for H_0).

(c) Now suppose the penalized estimator $\hat{\beta}$ is defined as

$$\begin{aligned}\hat{\beta} &= \arg\min_{b\in\mathbb{R}^{\prime}} Q_N(\beta) \\ &\equiv \arg\min_{b\in\mathbb{R}^{\prime}} \left(\left(\frac{1}{N}\sum_{i=1}^{N} |y_i - x_i' b| \right) + \frac{1}{2\sqrt{N}}(b-\delta_0)' A_0 (b-\delta_0) \right) \\ &\equiv \arg\min_{b\in\mathbb{R}^{\prime}} \left(S_N(b) + \frac{1}{2\sqrt{N}}(b-\delta_0)' A_0 (b-\delta_0) \right),\end{aligned}$$

i.e., the original penalty term is multiplied by $\sqrt{N}$. Under what additional conditions (if any) will $\hat{\beta}$ be consistent for β_0?

(d) Again assuming that

$$\sqrt{N}\frac{\partial^- Q_N(\hat{\beta})}{\partial b} = o_p(1),$$

where $\partial^- Q_N(b)/\partial b$ is the subgradient of $Q_N(b)$, and that

$$\sqrt{N}\frac{\partial^- S_N(\hat{\beta})}{\partial b} = \sqrt{N}\frac{\partial^- S_N(\beta_0)}{\partial b} + H_0\sqrt{N}(\hat{\beta}-\beta_0) + o_p(1),$$

derive the form of the asymptotic distribution of $\hat{\beta}$, assuming $\hat{\beta}$ is consistent for β_0.

[Q2] Let $\{R_i\}_{i=1}^{N}$ be a simple random sample of the 1×3 vector $R_i = (X_i, Y_i, Z_i)$. We assume that

$$Y_i = Z_i\beta_0 + U_i, \quad \mathbb{E}[U_i | X_i] = 0, \tag{1}$$

with $Z_i \in \{0,1\}$ binary and

$$\Pr(Z_i = 1 | X_i = x) = \Phi(x\gamma_0) \tag{2}$$

with $\Phi(\cdot)$ the cumulative distribution function of a standard normal. You may assume that other "standard" regularity conditions hold as well.

Consider the following two-step estimation procedure. First, apply maximum likelihood to the probit model (2), obtain $\hat{\gamma}$ and construct $\hat{Z}_i \stackrel{def}{\equiv} \Phi(X_i\hat{\gamma})$ for $i = 1,\ldots,N$. Second, compute the least squares fit of Y_i onto $\hat{Z}_i$:

$$\hat{\beta}_{TS} = \frac{\sum_{i=1}^{N} Y_i\hat{Z}_i}{\sum_{i=1}^{N}\hat{Z}_i^2}. \tag{3}$$

This procedure is analogous two-stage least squares, with the first stage based upon the probit instead of the linear probability model.

(a) Construct a moment function

$$\psi(R,\gamma,\beta) = \begin{bmatrix} \psi_1(R,\gamma) \\ \psi_2(R,\gamma,\beta) \end{bmatrix}$$

such that the corresponding method-of-moments estimate of (γ_0, β_0), i.e., the solution to

$$\sum_{i=1}^{N} \psi\left(R_i, \hat{\gamma}, \hat{\beta}_{TS}\right) = 0,$$

is identical to that of the two step procedure described above.

(b) Verify that your moment function is valid, in the sense that

$$\mathbb{E}\left[\psi\left(R, \gamma_0, \beta_0\right)\right] = 0.$$

(c) Consider the alternative, infeasible, one-step estimator $\hat{\beta}_{OS}$ which replaces $\hat{Z}_i$ in (3)with the true conditional probability $Z_{0i} \overset{def}{\equiv} \Phi\left(X_i\gamma_0\right)$. Compare the asymptotic variances of $\hat{\beta}_{TS}$ and $\hat{\beta}_{OS}$. Provide a characterization of the efficiency loss, if any, associated with having to estimate γ_0.

(d) Now assume that (2) is no longer valid (i.e., that the probit first stage is misspecified such that there is no γ_0 such that (2) holds for all $x \in \mathbb{X}$). For example it might be that

$$\Pr\left(Z_i = 1 | X_i = x\right) = \Phi\left(x\gamma_0 + x^2\delta_0\right).$$

Is the two-step estimator consistent under misspecification of the first stage (in general)? Explain.

(e) Consider the two-step instrumental variables estimator with $\psi_1(R, \gamma)$ as in part (a) above, but now

$$\psi_2\left(R_i, \gamma, \beta\right) = \left(Y_i - Z_i\beta\right)\Phi\left(X_i\gamma\right).$$

Does consistency of this estimator require both (1) and (2) or just the former or just the latter? Explain.

(f) Consider a second instrumental variables estimator with

$$\psi_2\left(R_i, \gamma, \beta\right) = \left(Y_i - Z_i\beta\right)X_i$$

and $\psi_1(R, \gamma)$ as in part (a) above. Assume that $\mathbb{V}\left(U_i | X_i = x\right) = \sigma^2$ for all $x \in \mathbb{X}$. Would you (generally) expect that the estimate of β_0 from part (e) to be more precisely, or less precisely, determined than the one based on the above moment function? Explain.

[Q3] Suppose $\{y_t : 1 \le t \le T\}$ is an observed strictly stationary time series generated by the (AR(1)) model

$$y_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i},$$

where $\phi \in (-1, 1)$ and $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0, 1)$.

As estimators of $\theta = \phi^2 \in [0, 1)$, consider

$$\hat{\theta} = \hat{\phi}^2, \qquad \hat{\phi} = \frac{\sum_{t=2}^{T} y_{t-1}y_t}{\sum_{t=2}^{T} y_{t-1}^2},$$

$$\tilde{\theta} = \frac{\sum_{t=3}^{T} y_{t-2}y_t}{\sum_{t=3}^{T} y_{t-2}^2},$$

and

$$\check{\theta} = \max(\tilde{\theta}, 0).$$

It can be shown that, for $k \in \{1, 2\}$,

$$\frac{1}{T} \sum_{t=k+1}^{T} y_{t-k}^2 \to_p E\left(y_{t-k}^2\right)$$

and

$$\frac{1}{\sqrt{T}} \sum_{t=k+1}^{T} y_{t-k}(y_t - \phi^k y_{t-k}) \to_d \mathcal{N}\left(0, \lim_{T\to\infty} Var\left[\frac{1}{\sqrt{T}} \sum_{t=k+1}^{T} y_{t-k}(y_t - \phi^k y_{t-k})\right]\right).$$

(a) Find the limiting distribution (after appropriate centering and rescaling) of $\hat{\theta}$.
(b) Find the limiting distribution of $\tilde{\theta}$.
(c) Find the limiting distribution of $\check{\theta}$.
(d) Rank the estimators from (a)-(c) in terms of (asymptotic) efficiency.

[Q4] Suppose $\{y_t : -1 \leq t \leq T\}$ is an observed strictly stationary time series generated by the (AR(1)) model

$$y_t = \varepsilon_t + \theta_0 \varepsilon_{t-1},$$

where $\theta_0 \in (-1, 1)$ and $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0, 1)$.

(a) Let $x_t = (y_t, y_{t-1}, y_{t-2})'$ and define the function

$$h(x_t, \theta) = \begin{bmatrix} y_{t-1}y_t - \theta \\ y_{t-2}y_t \end{bmatrix}.$$

Show that $\Theta = \{\theta_0\}$, where $\Theta = \{\theta : E[h(x_t, \theta)] = 0\}$.
Let

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

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

It can be shown that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} h(x_t, \theta_0) \to_d \mathcal{N}\left(0, \lim_{T\to\infty} Var\left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} h(x_t, \theta_0)\right]\right).$$

(b) It can be shown that

$$\sqrt{T}(\hat{\theta}_W - \theta_0) \to_d \mathcal{N}(0, \omega_W^2),$$

where ω_W^2 is some function of ϕ_0 and W. Verify this claim and express ω_W^2 in terms of θ_0 and W.

(c) Find W^*, a value of W for which ω_W^2 is minimal, and express $\omega_{W^*}^2$ in terms of θ_0.

(d) Propose a feasible estimator $\hat{\theta}$ (i.e., an estimator $\hat{\theta}$ that can be computed without knowledge of θ_0) satisfying

$$\sqrt{T}(\hat{\theta} - \theta_0) \to_d \mathcal{N}(0, \omega_{W^*}^2).$$