Department of Economics

UCLA

Second-Year Field Examination in Econometrics

Spring 2008

Instructions:

Solve all three parts I-III.

Use a separate bluebook for each part.

Solve all questions in each part.

You have 4 hours to complete the exam.

Calculators and other electronic devices are not allowed.

PART I (Based on Matzkin's course)

Problem 1:

Consider the model

$$y = \begin{cases} 1 & \text{if} \quad \beta x + m(z,\omega) \geq \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

where y , x, and z are observable, ω and ε are unobservable, x, ω, and ε are scalars, z is a $K-th$ dimensional vector, and m is an unknown function such that, for all values of z, m is strictly increasing in ω. Assume that x, z, ω and ε are mutually independent - the joint distribution of $(x, z, \omega, \varepsilon)$ equals the multiplication of the marginal distributions of x, z, ω, and ε. Assume also that the distributions of all random variables are such that the following result is true:

If W_1 and W_2 are independent random variables, and if the distributions of W_2 and of $W_1 + W_2$ are known, then the distribution of W_1 is known.

Answer the following questions:

(a) Suppose that the marginal distributions of ε, F_ε, and of ω, F_ω, are both known and strictly increasing. Let f_ε and f_ω denote, respectively, the densities. Suppose also that the support of x is R, the support of z is R^K, and that for a value $\overline{z}$ of z, it is known that

$$m(\overline{z},\omega) = \omega \quad \text{for all } \omega \text{ in } R$$

Define the unobservable variable w^* by $w^* = m(z,\omega)$, and the unobservable variable η by $\eta = \varepsilon - w^*$.

(a.1) Write down an expression for the probability of $y = 1$ given (x, z)

(a.2) Show that the parameter β is identified from the distribution of y given (x, z), when $z = \overline{z}$.

(a.3) Show that the distribution of η conditional on z is identified.

(a.4) Show that the function m is identified.

(b) Suppose next that both, F_ε and F_ω are unknown and strictly increasing, the support of x is R, the support of z is R^K, and for values $\widetilde{z}$ and $\overline{z}$ of z, it is known that

$$\begin{aligned} m(\widetilde{z},\omega) &= 0 \quad \text{for all } \omega \text{ in } R \\ m(\overline{z},\omega) &= \omega \quad \text{for all } \omega \text{ in } R \end{aligned}$$

PART II (Based on Vuong's course)

Let $(Y_i, X_i) \in \mathbb{R} \times \mathbb{R}^p$, $i = 1, \ldots, n$ be n iid observations of the random vector (Y, X), where X_i has density $f_X(\cdot)$.

1. Assume that $p = 1$ and suppose that $f_X(\cdot)$ is discontinuous at x (though continuous elsewhere). Specifically, $\lim_{\tilde{x}\uparrow x} f_X(\tilde{x}) = a_-$ while $\lim_{\tilde{x}\downarrow x} f_X(\tilde{x}) = a_+$ with $a_- \neq a_+$. Consider estimating $f_X(x)$ using a kernel estimator $\hat{f}_X(x)$ where $K(\cdot)$ is the uniform kernel on $[-1/2, +1/2]$. Give an expression for $\hat{f}_X(x)$ where h is the bandwidth. Give its expression and interpret it. Show that $\hat{f}_X(x)$ converges in mean square error to $f_X(x)$ as $h \downarrow 0$ and $nh \to \infty$ if and only if $a_+ + a_- = 2f_X(x)$.
Hint: Using the Mean Value Theorem on $[x - h/2, x]$ and $[x, x + h/2]$, show that $E[\hat{f}_X(x)] \to (1/2)(a_+ + a_-)$ as $h \downarrow 0$.

2. Assume that $p = 1$ and suppose that $f_X(\cdot)$ is continuous around x with $f_X(x) > 0$. Noting that $F_{Y|X}(y|x) = \mathrm{E}[\mathbb{I}(Y \leq y)|X = x]$, where $\mathbb{I}(\cdot)$ is the indicator of the event within parentheses, propose a kernel estimator $\hat{F}_{Y|X}(y|x)$ of $F_{Y|X}(y|x)$ using the uniform kernel on $[-1/2, +1/2]$ and interpret it. Show that $\hat{F}_{Y|X}(y|x) \to F_{Y|X}(y|x)$ in probability as the bandwidth $h \downarrow 0$ and $nh \to \infty$.
Hint: Writing $\hat{F}_{Y|X}(y|x)$ as a fraction $\hat{\phi}(y, x)/\hat{f}_X(x)$, establish the convergence in quadratic mean of the numerator, while using question **1** for $\hat{f}_X(x)$. Assume that $\partial F_{YX}(y, x)/\partial x$ exists and is continuous around x, where $F_{YX}(y, x) = \int_{-\infty}^{x} F_{Y|X}(y|\tilde{x}) f_X(\tilde{x}) d\tilde{x}$ is the joint cdf of (Y, X) at (y, x).

3. Assume that $p = 2$ so that $X = (X_1, X_2) \in \mathbb{R}^2$. For a given value y, consider testing whether $F(y|X_1, X_2)$ depends on X_2, i.e. testing $H_0 : F_{Y|X}(y|X) = F_{Y|X_1}(y|X_1)$ vs. $H_1 : F_{Y_X}(y|X) \neq F_{Y|X_1}(y|X_1)$. Show that the conditional moment restriction H_0 holds *if and only if* the unconditional moment restriction

$$H_0' : \mathrm{E}\Big\{[\mathbb{I}(Y \leq y) - F_{Y|X_1}(y|X_1)]\Psi(X)]\Big\} = 0$$

holds, where $\Psi(X) = \omega(X)[F_{Y|X}(y|X) - F_{Y|X_1}(y|X_1)]$ and $\omega(X) > 0$. Taking $\omega(X) = f_{X_1}^2(X_1) f_X(X)$, propose a test statistic that can be used to test H_0. In particular, why can this choice of $\omega(\cdot)$ simplify the study of the properties of the test statistic? If Y is a dummy variable taking the values 0 or 1, what is a natural choice for y?
Note: Define precisely each term in your test statistic. You do not have to derive the asymptotic properties of your test statistic.

PART III (Based on Hahn's course)

1. Assume fixed T asymptotics for this question. Consider the panel version of simultaneous equations model

$$\begin{aligned} y_{it} &= \alpha_i + x_{it}\theta + \varepsilon_{it} \\ x_{it} &= \gamma_i + z_{it}\pi + v_{it}, \quad (t = 1, ..., T; i = 1, \ldots, n) \end{aligned}$$

For simplicity, we will assume that every variable is a scalar. We will consider the 2SLS estimation of θ by the following steps:

Step 1 Estimate γ_i and π by OLS. Call the estimators $\widehat{\gamma}_i$ and $\widehat{\pi}$.

Step 2 Estimate α_i and θ by OLS regression of the first equation, replacing x_{it} by $\widehat{x}_{it} = \widehat{\gamma}_i + z_{it}\widehat{\pi}$. The resultant estimator for θ will be called $\widehat{\theta}$.

Show that $\widehat{\pi}$ and $\widehat{\theta}$ can be understood to be method of moments estimator from the moment equation

$$0 = E\left[\sum_{t=1}^{T} \widetilde{z}_{it}\left(\widetilde{x}_{it} - \widetilde{z}_{it}\pi\right)\right]$$

$$0 = E\left[\sum_{t=1}^{T} \left(\widetilde{z}_{it}\pi\right)\left(\widetilde{y}_{it} - \left(\widetilde{z}_{it}\pi\right)\theta\right)\right]$$

Are these moment equations valid at the true values of the parameters π and θ? Do you conclude that the 2SLS is consistent under the fixed T asymptotics? HINT: From the derivation of the fixed effects estimator, you may recall that $\widehat{\pi}$ can be obtained by regressing $\widetilde{x}_{it} = x_{it} - \overline{x}_i$ on $\widetilde{z}_{it} = z_{it} - \overline{z}_i$, or

$$0 = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} \widetilde{z}_{it}\left(\widetilde{x}_{it} - \widetilde{z}_{it}\widehat{\pi}\right)$$

From the same perspective, we can view that $\widehat{\theta}$ is obtained by regressing $\widetilde{y}_{it} = y_{it} - \overline{y}_i$ on

$$\widetilde{\widehat{x}}_{it} = \widehat{x}_{it} - \overline{\widehat{x}}_i = \widehat{\gamma}_i + z_{it}\widehat{\pi} - \frac{1}{T}\sum_{s=1}^{T}\left(\widehat{\gamma}_i + z_{is}\widehat{\pi}\right) = \widetilde{z}_{it}\widehat{\pi}$$

2. Consider the simultaneous equations model

$$\begin{aligned} y_i &= \theta_0 x_i + \varepsilon_i \\ x_i &= z_i'\pi + v_i \qquad (i = 1, ..., n) \end{aligned}$$

where we assume that z_i are non-stochastic. We also assume that

$$\frac{1}{n}Z' \underset{n\times q}{Z} = \frac{1}{n}\sum_i z_i z_i'$$

is <u>fixed</u> at Υ as $n \to \infty$. Finally, we assume that (ε_i, v_i) is bivariate normal. Note that the 2SLS $\widehat{\beta}$ is such that

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)=\frac{\left(\frac{1}{n}X'Z\right)\Upsilon^{-1}\left(\frac{1}{\sqrt{n}}Z'\varepsilon\right)}{\left(\frac{1}{n}X'Z\right)\Upsilon^{-1}\left(\frac{1}{n}Z'X\right)}$$

Let

$$\omega_1 \equiv \frac{1}{\sqrt{n}}Z'\varepsilon, \quad \omega_2 \equiv \frac{1}{\sqrt{n}}Z'v$$

(a) Show that

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_\varepsilon^2\Upsilon & \sigma_{\varepsilon v}\Upsilon \\ \sigma_{\varepsilon v}\Upsilon & \sigma_v^2\Upsilon \end{bmatrix}\right)$$

(b) Show that

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)=\frac{\pi'\omega_1+\frac{1}{\sqrt{n}}\omega_2'\Upsilon^{-1}\omega_1}{\pi'\Upsilon\pi+\frac{2}{\sqrt{n}}\omega_2'\pi+\frac{1}{n}\omega_2'\Upsilon^{-1}\omega_2}$$

(c) Show that

$$\sqrt{n}\left(\widehat{\beta}-\beta\right)=\frac{\pi'\omega_1}{\pi'\Upsilon\pi}+\frac{1}{\sqrt{n}}\left(\frac{\omega_2'\Upsilon^{-1}\omega_1}{\pi'\Upsilon\pi}-\frac{2\left(\pi'\omega_1\right)\left(\omega_2'\pi\right)}{\left(\pi'\Upsilon\pi\right)^2}\right)+o_p\left(\frac{1}{\sqrt{n}}\right)$$

(d) Show that

$$E\left[\frac{\pi'\omega_1}{\pi'\Upsilon\pi}+\frac{1}{\sqrt{n}}\left(\frac{\omega_2'\Upsilon^{-1}\omega_1}{\pi'\Upsilon\pi}-\frac{2\left(\pi'\omega_1\right)\left(\omega_2'\pi\right)}{\left(\pi'\Upsilon\pi\right)^2}\right)\right]=\frac{q-2}{\pi'\Upsilon\pi}\sigma_{\varepsilon v}$$

Interpret this equality in relation to the finite sample bias of 2SLS.