Spring 2004
Department of Economics
UCLA

ECONOMETRICS FIELD EXAMINATION

Instructions: This is a 4 hour closed book/closed notes exam. There are FOUR parts in the exam. Answer ALL questions of ONLY TWO parts of your choice. Use a separate bluebook for each part. GOOD LUCK!

PART I

Question 1:

Suppose that

$$y_i = x_i\beta + u_i$$

and

$$E\left[u_i | x_i\right] = 0$$

We assume that $E\left[u_i^2 | z_i\right]$ is fixed at σ^2. Consider the two-step GMM estimator $\widehat{\beta}_{GMM}$ that solves

$$\min_b \left(\frac{1}{n}\sum_{i=1}^n z_i\left(y_i - x_i b\right)\right)' \widehat{A}^{-1} \left(\frac{1}{n}\sum_{i=1}^n z_i\left(y_i - x_i b\right)\right)$$

where

$$z_i = \begin{bmatrix} x_i \\ x_i^2 \end{bmatrix}$$

and

$$\widehat{A} \equiv \frac{1}{n}\sum_{i=1}^n \left(y_i - x_i\widehat{\beta}_{OLS}\right)^2 z_i z_i'$$

Derive the asymptotic distributions of $\sqrt{n}\left(\widehat{\beta}_{GMM} - \beta\right)$ and $\sqrt{n}\left(\widehat{\beta}_{OLS} - \beta\right)$. Show that the asymptotic variance of $\sqrt{n}\left(\widehat{\beta}_{GMM} - \beta\right)$ is equal to that of $\sqrt{n}\left(\widehat{\beta}_{OLS} - \beta\right)$.

Question 2:

Consider the data generated by variable probability sampling (Wooldridge). Below is the description of the sampling scheme which we discussed in class.

- The population is first partitioned into J (nonoverlapping and exhaustive) groups, $\mathcal{W}_1, \ldots, \mathcal{W}_J$.
- We draw an observation $w_i = (y_i, x_i)'$ at random from the population. $(i = 1, \ldots, N)$ For simplicity, we will assume that $\dim(x_i) = 1$.

- If w_i is in stratum j, flip a coin with probability p_j of turning up heads. Let $h_{ij} = 1$ if the coin turns up heads and zero otherwise.

- We keep the observation i if $h_{ij} = 1$; otherwise, omit it from the sample.

- Let s_{ij} denote the indicator of the stratum that w_i belongs to. For example, $s_{i1} = 1$ if $w_i \in \mathcal{W}_1$. Let $r_{ij} \equiv h_{ij} s_{ij}$. (By definition, $r_{ij} = 1$ for at most one j. If $h_{ij} = 1$, then $r_{ij} = s_{ij}$. If $r_{ij} = 0$ for all j, then the random draw w_i does not appear in the sample.)

- Let $q_i \equiv \frac{r_{ij}}{p_j}$ for such j, and call it the "sampling weights". We will assume that the sampling weight for each is reported along with w_i in the sample.

- Let N_j denote the number of observations falling into stratum j. Let $N_0 = N_1 + \cdots + N_J$. (Note that $N_0 \neq N$ in general under the variable probability sampling.)

1. Now, suppose that

$$y_i = x_i \cdot \beta + u_i \qquad i = 1, 2, \ldots$$

where $E[x_i u_i] = 0$. Without too much loss of generality, we may assume that the last $N - N_0$ observations are such that $r_{ij} = 0$ for all j. Let

$$\widehat{\beta}_w \equiv \min_b \sum_{i=1}^{N_0} q_i (y_i - x_i \cdot b)^2$$

(a) Show that

$$\widehat{\beta}_w = \min_b \sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} (y_i - x_i \cdot b)^2$$

and

$$\begin{aligned} \widehat{\beta}_w &= \left(\sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i^2 \right)^{-1} \left(\sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i y_i \right) \\ &= \beta + \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i^2 \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i u_i \right) \end{aligned}$$

(b) Show that

$$0 = E\left[\sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i u_i \right]$$

and

$$\widehat{\beta}_w = \beta + o_p(1)$$

(c) Show that

$$\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{J} \frac{r_{ij}}{p_j} x_i^2 = E\left[x_i^2\right] + o_p(1)$$

(d) Show that

$$E\left[\left(\sum_{j=1}^{J}\frac{r_{ij}}{p_j}x_iu_i\right)^2\right]=\sum_{j=1}^{J}\frac{1}{p_j}E\left[s_{ij}x_i^2u_i^2\right]$$

and

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{j=1}^{J}\frac{r_{ij}}{p_j}x_iu_i\to N\left(0,\sum_{j=1}^{J}\frac{1}{p_j}E\left[s_{ij}x_i^2u_i^2\right]\right)$$

(e) From previous questions, we can conclude that

$$\sqrt{N}\left(\widehat{\beta}_w-\beta\right)\to N\left(0,\frac{\sum_{j=1}^{J}\frac{1}{p_j}E\left[s_{ij}x_i^2u_i^2\right]}{\left(E\left[x_i^2\right]\right)^2}\right)$$

We would like to obtain a 95% confidence interval for β that does not depend on N. For this purpose, we will assume that u_i is observed. (Clearly a nonsense, but a useful assumption to simplify this question.) Show that

$$\frac{1}{N}\sum_{i=1}^{N_0}q_i^2x_i^2u_i^2=\sum_{j=1}^{J}\frac{1}{p_j}E\left[s_{ij}x_ix_i'u_i^2\right]+o_p(1)$$

and

$$\frac{1}{N}\sum_{i=1}^{N_0}q_ix_i^2=E\left[x_i^2\right]+o_p(1)$$

Based on these observations, suggest a 95% confidence interval for β that does not depend on N.

PART II

Question 1—Dynamic Programming:

Consider a three-period model in which an individual makes decisions about three variables: consumption c_t, leisure l_t, and whether or not to attend school d_t, at each year, for $t = 1, 2, 3$. Note that while the first two variables are continuous choice variables, d_t is a binary choice variable that takes the value 1 if an individual decides to attend school at year t, and 0 otherwise.

Each period utility is given by

$$u_t(c_t, l_t) = u(c_t, l_t) = \kappa c_t^{\gamma_1} l_t^{\gamma_2},$$

Each individual gets a stream of unearned income. In each period that amount is given by i_t. If in addition the individual works then he/she is paid an hourly wage w_t, which is a function of the number of years of education. Specifically, we have

$$\log w_t = \alpha_0 + \alpha_1 ed_t + \alpha_2 ed_t^2 + \varepsilon_t,$$

where *ed* denoted the education level and ε_t is an idiosyncratic shock, uncorrelated with *ed*. The individual starts with an endowment of A_0, and earned an interest rate of r_t for any unused monetary resources carried over from period t to period $t+1$. Assume that the interest rates for the three periods are known in advance.

For the questions specified below, if you think that some needed information has been omitted, please make assumption about that needed information.

1. Define the state vector, say z_t.
2. Specify all the necessary budget constraints, regarding money and time.
3. Write the value function at each period for the individual decision maker.
4. Provide the full list of parameters that need to be estimated. Let the true parameter vector be denoted by ϕ_0.
5. Provide details about the data that one would need in order to be able to estimate the parameters of the above model.
6. Provide detailed information about the method by which you propose to estimate the parameter vector ϕ_0. You should answer this question as if you were giving instructions to a professional programmer who knows nothing about economics. (Keep in mind that there are several alternative ways for obtaining an estimate for ϕ_0.)
7. What is the asymptotic distribution for the parameter vector estimate, say $\widehat{\theta}$, obtained following the procedure suggested in (6.)? Provide brief justifications for all your claims.

Question 2—Quantile Regression:

Consider the quantile regression model. That is,

$$y_i = x_i'\beta_\theta + u_{\theta i}, \tag{2.1}$$

where the θth conditional quantile of y_i, conditional on x_i, satisfies

$$Q_\theta\left(y_i|x_i\right) = 0, \tag{2.2}$$

for $i = 1, ..., n$.

1. Suppose you are asked to use the quantile regression model in order to provide a discrete approximation for the continuous *conditional* log-wage distribution (such as in question 1 above), that is, the conditional distribution of the dependent variable conditional on a vector of observed characteristics vector x_i. Describe how to use the quantile regression to provide such an approximation.

2. Suppose you estimate three quantile regression for θ_1, θ_2, and θ_3. How would you test the hypothesis that the slope coefficients in the corresponding parameter vectors, i.e., β_{θ_1}, β_{θ_2}, and β_{θ_3}, are all the same.

3. Suppose that

$$Q_\theta\left(u_{\theta i}|x_i\right) \neq 0,$$

but there exists a vector z_i, with $\dim\left(z_i\right) = \dim\left(x_i\right)$, such that

$$Q_\theta\left(u_{\theta i}|z_i\right) = 0.$$

(a) Show that the estimator $\widehat{\beta}_\theta$, given by

$$\widehat{\beta}_\theta = \arg\min_\beta \frac{1}{n}\sum_{i=1}^{n}\rho_\theta\left(y_i - x_i'\beta\right),$$

is not a consistent estimator for the population parameter β_θ. The function $\rho_\theta\left(\cdot\right)$ is the check function.

(b) Under what conditions, if at all, would the estimator provided by a solution to

$$\frac{1}{n}\sum_{i=1}^{n}\left(\theta - 1/2 + 1/2\text{sgn}\left(y_i - x_i'\widehat{\beta}_{IV}\right)\right)z_i = 0,$$

be consistent estimator for β_θ? Explain briefly. The function sgn$(\cdot)$ is simply the sign function, that is, sgn$(\lambda) = 1$ if $\lambda \geq 0$, and sgn$(\lambda) = -1$ otherwise.

PART III

Question 1:

Consider the static linear panel data model

$$y_{it} = x_{it}\beta_0 + \alpha_i + \varepsilon_{it}$$

which in matrix form can be written as

$$Y = X\beta_0 + D\alpha + \varepsilon$$

where as usual Y and ε are $NT \times 1$ vectors, X is an $NT \times K$ matrix, a is a $N \times 1$ vector, and $D = I_N \otimes e_T$ with e_T the $T-$ dimensional unit vector. We want to test for random versus fixed effects. Assume that

$$E(\varepsilon|X,\alpha) = 0 \qquad \text{(A1)}$$

$$E(\varepsilon\varepsilon'|X,\alpha) = V(\varepsilon|X,\alpha) = \sigma_\varepsilon^2 I_{NT} \qquad \text{(A2)}$$

while under the null the following two assumptions hold:

$$E(\alpha|X) = E(\alpha) = 0 \qquad \text{(A3)}$$

$$E(\alpha\alpha'|X) = V(\alpha|X) = V(\alpha) = \sigma_\alpha^2 I_N \qquad \text{(A4)}$$

(a) Derive the asymptotic distribution of the Hausman test which is based on the difference between the random effects and fixed effects estimators of β_0.

(b) Show that the test in (a) is numerically equal to the Hausman test which is based on the difference between the between effects and fixed effects estimators of β_0.

(c) Show that the test in (b) is nothing but the Wald test of the hypothesis that $\gamma = 0$ in the following augmented model:

$$Y^* = X^*\beta + \tilde{X}\gamma + w$$

where X^* and Y^* are the GLS transformed X and Y, i.e. $X^* = \Omega^{-1/2}X$ and $Y^* = \Omega^{-1/2}Y$, and $\tilde{X}$ is the within transformed X, i.e. $\tilde{X} = QX$.

Question 2:

(a) Define the maximum score estimator for the cross sectional binary choice model

$$y_i = 1\{x_i\beta_0 + \varepsilon_i > 0\} \qquad i = 1, 2, \ldots, N$$

Give the conditions under which it is consistent and the intuition behind the estimator.

(b) Do the same as in part (a) for the static fixed effects panel data version of the model

$$y_{it} = 1\{x_{it}\beta_0 + \alpha_i + \varepsilon_{it} > 0\} \qquad i = 1, 2, \ldots, N; \qquad t = 1, 2, ..., T$$

where $T \geq 2$ is fixed.

(c) Consider now the AR(1) fixed effects version of the model

$$y_{it} = 1\{y_{it-1}\gamma_0 + \alpha_i + \varepsilon_{it} > 0\} \qquad i = 1, 2, \ldots, N; \qquad t = 1, 2, ..., T$$

Can the maximum score estimator be used to estimate γ_0 and how? What are sufficient conditions?

(d) Describe the random effects approach for estimating the models in (b) and (c). Discuss the advantages and disadvantages of the approach as opposed to the fixed effects approach.

PART IV

Question 1: True/Questionable/False. In this question, just stating true, questionable or false will not give any points. It is the explanation that counts.

(i) If one regresses a random walk $y_{1,t}$ on its own lagged observation $y_{1,t-1}$, on an independent random walk $y_{2,t}$, and on its lagged observation $y_{2,t-1}$

$$y_{1,t} = \widehat{\alpha} y_{1,t-1} + \widehat{\beta} y_{2,t} + \widehat{\gamma} y_{2,t-1}$$

then the theory of spurious regression implies that the OLS estimator $\widehat{\beta}$ of β is inconsistent.
(ii) For consistent estimation of the AR-parameter ρ of a causal AR(1) process y_t with mean α, we can use OLS in the regression $y_t = \alpha + \rho y_{t-1} + \varepsilon_t$ even if the error term ε_t is serially correlated.
(iii) Denote by " $\Rightarrow$ " convergence in law for random functions on the unit interval. Assume $S_T(\cdot) \Rightarrow 0$ as $T \to \infty$ and let $r \in [0,1]$. Then $S_T(r) \to_p 0$.
(iv) The optimal linear forecast of a time series y_t based on some of its lagged observations does not depend on third and higher moments of the series y_t *only if* the time series is Gaussian.
(v) With independent data (u_t, x_t), kernel HAC estimation of the variance covariance matrix of the OLS estimator $\widehat{\beta}$ in the regression $y_t = x_t\beta + u_t$ is consistent *even if* the bandwidth S_T does not grow to infinity as the sample size T goes to infinity.
(vi) The optimal *linear* forecast of a time series y_t based on some other variables X_{t-1} is always unbiased.

Question 2:
(i) Give detailed definitions of the concepts of ergodicity and α-mixing for a time series process y_t.
(ii) State strong law of large numbers (SLLN) for ergodic and α-mixing processes.
(iii) Compare and discuss the assumptions needed for the statements in (ii). Explain in what sense the assumptions needed for a SLLN under ergodicity are more stringent than the ones for a SLLN under α-mixing. Explain in what sense the assumptions needed for a SLLN under α-mixing are more stringent than the ones for a SLLN under ergodicity.

Question 3:
(i) Prove directly (without stating any theorem) that a random walk is not weakly stationary.
(ii) Prove in detail that if a causal AR(1) process is observed every other time period then the resulting process again is a causal AR(1) process.

Question 4:
For $t = 1, ..., T$, let

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix},$$

where $\delta_2 \neq 0$. Let $u_t := (u_{1t}, u_{2t})'$, and suppose that $u_t = \Psi(L)\varepsilon_t$ for ε_t an i.i.d. (2×1) vector with mean zero, variance PP', and finite fourth moments. Assume further that $\{s\Psi_s\}_{s=0}^{\infty}$ is absolutely summable and that $\Psi(1)P$ is nonsingular. For $i = 1, 2$, define $\xi_{it} := \sum_{s=1}^{t} u_{is}$, and $\gamma_0 := \delta_1/\delta_2$.

(i) Using Donsker's theorem and the Beveridge-Nelson decomposition, outline how to derive the asymptotic distribution of $T^{-3/2}\sum_{t=1}^{T}(\xi_{1t}, \xi_{2t})'$.

(ii) Show that the OLS estimates of

$$y_{1t} = \alpha + \gamma y_{2t} + u_t$$

satisfy

$$\left[\begin{array}{c} T^{-1/2}\widehat{\alpha}_T \\ T^{1/2}(\widehat{\gamma}_T - \gamma_0) \end{array}\right] - \left[\begin{array}{cc} 1 & \delta_2/2 \\ \delta_2/2 & \delta_2^2/3 \end{array}\right]^{-1} \left[\begin{array}{c} T^{-3/2}\sum\limits_{t=1}^{T}(\xi_{1t} - \gamma_0\xi_{2t}) \\ T^{-5/2}\sum\limits_{t=1}^{T}\delta_2 t(\xi_{1t} - \gamma_0\xi_{2t}) \end{array}\right] \to_p 0.$$

Conclude that $\widehat{\alpha}_T$ and $\widehat{\gamma}_T$ have the same asymptotic distribution as the coefficients from a regression of $(\xi_{1t} - \gamma_0\xi_{2t})$ on a constant and δ_2 times a time trend:

$$(\xi_{1t} - \gamma_0\xi_{2t}) = \alpha + \gamma(\delta_2 t) + u_t.$$

(iii) Show that first differences of the OLS residuals converge to

$$\Delta\widehat{u}_t - (u_{1t} - \gamma_0 u_{2t}) \to_p 0.$$