Econometrics Preliminary Exam

Agricultural and Resource Economics, UC Davis

July 2, 2015

There are **FOUR** questions. Answer each part of each question. All questions are weighted equally. Within each question, each part will receive equal weight in grading.

I. U is a uniform random variable observed on the $(0, 1)$ interval and X is a random variable defined as the sixth-root of U, i.e., $X = U^{1/6}$. The probability density function for a $Uniform(a, b)$ random variable is

$$f(U) = \begin{cases} \frac{1}{b-a} & \text{if } a \leqslant U \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the cumulative distribution function and the probability density function for the new random variable X.

(b) Suppose that X_1, X_2, and X_3 are *iid* random variables generated by drawing three *iid Uniform*$(0, 1)$ random variables and finding their sixth-roots, as specified above. Find the probability that exactly two of the three X_i random variables are less than 0.5.

(c) Suppose now that you know only that $X = U^{\alpha}$. Your experiment will consist of observing just one observation for X. Derive a likelihood-ratio test of the hypothesis that $\alpha = 1$ against the alternative hypothesis that $\alpha = 0.5$. Give exact values for the probability of a Type-I error, the power of your test, and the probability of a Type-II error.

II. Let $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$, where $u_i | X_i \overset{i.i.d.}{\sim} N(0, 1)$.

(a) Propose an OLS, a maximum likelihood and a method of moments estimator for $\beta = (\beta_1, \beta_2)'$. For each estimator, clearly define the objective function that defines it and solve for a closed-form solution if it exists.

(b) Are any of the estimators you proposed in (a) efficient? Give your reasoning.

(c) Now assume that the above model is not correctly specified and that instead $y_i = \gamma_1 x_{1i} \quad x_{2i}^{\gamma_2} \quad \epsilon_i$, where $E[\epsilon_i \; x_{i1}, x_{i2}] = 0$. If you would still like to use the linear model for estimation, say using OLS regression of y_i on x_{1i} and x_{2i}, can you identify the conditional expectation $E[y_i \; x_{i1}, x_{i2}]$? Justify your answer.

(d) Now propose an estimator of $\gamma = (\gamma_1, \gamma_2)'$. Clearly define its objective function and give a closed-form solution for the estimator if it exists. If there is no closed-form solution, explain why.

(e) Give conditions for consistency of the estimator you proposed in (d). Provide primitive conditions wherever possible.

(f) Now suppose that $E[\epsilon_i \; x_{1i}, x_{2i}] \neq 0$, but you have instruments $z_i = (z_{1i}, z_{2i}, z_{3i})'$ that satisfy $E[z_i \epsilon_i] = 0$. Propose an efficient estimator of γ_1 and γ_2.

(g) One of your friends proposes that you run a first stage regression of x_i on z_i to plug in the fitted value of x_i, $\hat{x}_i$ into the equation $y_i = \gamma_1 x_{1i} \quad x_{2i}^{\gamma_2} \quad \epsilon_i$ in lieu of x_i. Would this procedure yield a consistent estimator of γ_1 and γ_2?

III. Consider the following panel of individuals (i) from different states (s) across time (t), where $y_{ist} = x'_{ist}\beta \quad \lambda_i \quad \gamma_s \quad u_{ist}$. Let $x_{is} = (x_{is1}, x_{is2}, \ldots, x_{isT})$ and $dim(x_{ist}) = dim(\beta) = k$. Suppose $E[u_{ist} \; x_{is}, \lambda_i, \gamma_s] = 0$, but $E[\lambda_i \; x_{is}] \neq 0$ and $E[\gamma_s \; x_{is}] \neq 0$.

(a) To estimate β consistently, do you need to include individual and state fixed effects in the regression? Give a formal <u>and</u> an intuitive explanation for your answer.

(b) How would your answer in (a) change if $y_{ist} = x'_{ist}\beta \quad \lambda_i \quad \gamma_{st} \quad u_{ist}$. Explain your reasoning.

(c) Propose a transformation of the model in (b) to estimate β without estimating λ_i and γ_{pt}. How would the asymptotic distribution of the OLS estimator using this transformation change if $Var(\lambda_i)$ changes?

Note: No need to derive the relevant asymptotic distribution.

(d) Now suppose you are only observing individuals in the same state, so your model is $y_{it} = x'_{it}\beta \quad \lambda_i \quad u_{it}$, where $E[u_{it} \; X_i, \lambda_i] = 0$ and $E[\lambda_i \; X_i] \neq 0$. Propose an estimator and show its consistency and asymptotic normality. Give an expression of its asymptotic variance.

(e) Propose a test for the joint significance of all k elements of β, i.e. $H_0 : \beta = 0$.

IV. Consider the following model for the supply of widgets:

$$q_t = \beta_0 \quad \beta_1 p_t \quad \beta_2 x_t \quad \epsilon_{1t}$$

where q_t denotes the log of quantity, p_t denotes the log of price and x_t denotes the log of energy prices. You have time series data on q_t, p_t, and x_t. The parameter β_1 can be interpreted as the price elasticity of supply. Suppose that energy prices are determined exogenously to the widget market.

Suppose the demand for widgets is perfectly elastic (i.e., horizontal demand curve), and that price is determined by the equation

$$p_t = \alpha p_{t \ 1} \quad \epsilon_{2t}$$

where $\ \alpha\ < 1$. Assume that ϵ_{1t} and ϵ_{2t} are *iid* with mean zero and uncorrelated with each other at all leads and lags.

(a) Show that OLS produces a consistent estimate of β_0, β_1, and β_2. State any assumptions you make.

(b) Suppose you hypothesize that widget manufacturers choose their production level q_t in period $t-1$ based on the expected period t price, i.e., you hypothesize that the correct supply model is

$$q_t = \beta_0 \quad \beta_1 E[p_t \ \Im_{t \ 1}] \quad \beta_2 x_t \quad \epsilon_{1t}$$

where $\Im_{t \ 1}$ denotes information available at time $t-1$. Under this hypothesis, would OLS regression of q_t on a constant, p_t, and x_t provide consistent estimates of β_0, β_1, and β_2? Justify your answer with a proof.

(c) Continuing with the hypothesis in (b), would an IV estimator that uses $p_{t \ 1}$ to instrument for p_t provide consistent estimates of β_0, β_1, and β_2? Justify your answer with a proof.

(d) Continuing with the hypothesis in (b), would an IV estimator that uses $p_{t \ 1}$ and $p_{t \ 2}$ to instrument for p_t provide more efficient estimates of β_0, β_1, and β_2 than the IV estimator in (c)? Justify your answer. A mathematical proof is not necessary.

(e) Write down the statistic you would use to test the hypothesized model in (b) against the original supply model at the beginning of this question. State the asymptotic null distribution of your statistic.

(f) Suppose $\alpha = 0$. Describe the implications of this fact for your IV estimator in (c).

(g) Suppose α is close to zero. Describe the implications of this fact for the IV estimator in (c). How would you determine whether α is far enough away from zero for you to have confidence in your IV estimates?

Notation. θ_0, Θ, y_i, x_i, w_i, $s(y_i, x_i; \theta)$, $H(y_i, x_i; \theta)$ and $h(y_i, w_i; \theta)$ pertain to the objects defined in the 240B lecture notes.

Assumption (Uniform Law of Large Numbers $\{\sup_{\theta\in\Theta} |\sum_{i=1}^n f(y_i, w_i; \theta)/n - E[f(y_i, w_i; \theta)]| \xrightarrow{p} 0\}$**)**

(i) *(i.i.d.)* $\{y_i, w_i\}_{i=1}^n$ is an i.i.d. sequence of random variables;

(ii) *(Compactness)* Θ is compact;

(iii) *(Continuity)* $f(y_i, w_i; \theta)$ is continuous in θ for all $(y_i, w_i')'$;

(iv) *(Measurability)* $f(y_i, w_i; \theta)$ is measurable in $(y_i, w_i')'$ for all $\theta \in \Theta$;

(v) *(Dominance)* There exists a dominating function $d(y_i, w_i)$ such that $|f(y_i, w_i; \theta)| \leqslant d(y_i, w_i)$ for all $\theta \in \Theta$ and $E[d(y_i, w_i)] < \infty$.

Assumption (Consistency of Sample Average of Hessian for M-Estimators)

(i) Each element of $H(y_i, x_i; \theta)$ is bounded in absolute value by a function $b(y_i, x_i)$, where $E[b(y_i, x_i)] < \infty$;

(ii) $A_0 = -E[H(y_i, x_i; \theta_0)]$ is positive definite.

Assumption (Asymptotic Normality of Sample Average of Score for M-Estimators)

(i) $E[s(y_i, x_i; \theta_0)] = 0$;

(ii) each element in $s(y_i, x_i; \theta_0)$ has finite second moment.

Formula for the score statistic

$$S \equiv \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n s_i(\hat{\theta}_R)\right)' A_{nR}^{-1} C_{nR}' \left\{\widehat{Avar}\left(C_{nR} A_{nR}^{-1} \sum_{i=1}^n s_i(\hat{\theta}_R)/\sqrt{n}\right)\right\}^{-1} C_{nR} A_{nR}^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^n s_i(\hat{\theta}_R)\right)$$

GMM Expression for the Sampling Error

$$\sqrt{n}(\hat{\theta} - \theta_0) = -\left(\mathcal{H}_0' W \frac{1}{n}\sum_{i=1}^n \left.\frac{\partial h(y_i, w_i; \theta)}{\partial \theta}\right|_{\theta=\theta*}\right)^{-1} \mathcal{H}_0' W \frac{1}{\sqrt{n}} \sum_{i=1}^n h(y_i, w_i; \theta_0) + o_p(1)$$

where

$$\mathcal{H}_0 = E\left[\left.\frac{\partial h(y_i, w_i; \theta)}{\partial \theta'}\right|_{\theta=\theta_0}\right]$$