June 25, 2002

UCLA Department of Economics

Econometrics Field Comprehensive Examination
Spring 2002

Answer **six** questions as follows: Questions 1 & 2 in Parts I and II and question 1 in Part III are mandatory; then choose one more to answer from question 3 in Part I or II or question 2 in Part III. Each question has equal weight. **4 hours**

PART I

1. A distribution that is used for duration data, e.g., unemployment durations is the exponential distribution. Suppose that conditional on education (X) the duration of unemployement spells (Y) has an exponential distribution with parameter $\beta_0 + \beta_1 X$.

   (a) Describe how you would estimate $\beta = (\beta_0, \beta_1)$ by maximum likelihood given a random sample of size N.

   (b) What is the variance of the maximum likelihood estimator?

   (c) Show that the least squares estimator for the slope coefficient of the regression of Y on a constant and X is consistent for β_1

   (d) How does the variance of the least squares estimator compare to that of the maximum likelihood estimator?

   (e) How can you modify the least squares estimator to make it efficient?

2. Consider the following linear model

$$E[Y|X] = X\beta.$$

Suppose we consider estimating β using a generalized method of moments framework with two moment functions

$$\psi_1(Y, X, \beta) = X \cdot (Y - X\beta),$$

and

$$\psi_2(Y, X, \beta) = Y \cdot X - 2,$$

where Y, X and β are all scalars.

(a) Describe the optimal gmm estimator for β.

(b) Describe the empirical likelihood estimator for β.

(c) Compare the variance of the optimal gmm estimator using both moments with the variance of the gmm estimator using only the first moment.

(d) Describe how you would estimate the variance using bootstrapping.

3. Suppose you are interested in estimating the average effect of a job training program using observational data.

(a) Suppose you have detailed data on the background of the participants in the program and nonparticipants, and you are willing to make the unconfoundedness assumption / selection on observables. Describe two ways how you could implement such a strategy.

(b) Suppose you have an instrumental variable that only affects the likelihood of participating in the program, but not the outcomes directly. Describe how you could use this to estimate the average effect of the program, and discuss any differences with the previous strategies.

PART II

1. (a) Answer the following true/false questions. No points without a correct **brief** explanation.

   i. Two random variables, $Y \sim N(0,1)$, $Z \sim N(0,1)$, when uncorrelated, are also independent.

   ii. Let I denote the indicator function and let $\{U_t\}$ be i.i.d.. Then the process $\{Z_t\}$ with

$$Z_t = U_t\Big(I(t \text{ odd})\sqrt{|Z_{t-1}|/E|Z_{t-1}|} + I(t \text{ even})\sqrt{Z_{t-1}^2/EZ_{t-1}^2}\Big) \tag{1}$$

   is covariance–stationary.

   Let $\{X_t\}$ be a time series of length T.

   (b) Assume for the purpose of this subquestion that $\{X_t\}$ is strictly stationary. Explain **briefly** why the sample mean $\bar{X}$ does not necessarily converge to the population mean $\mu = EX_1$ like it (generally) does in a cross section.

   (c) Let $\mathcal{G}^t$ be the σ–algebra[1] generated by $\{X_t\}$ up to period t. Suppose that

$$E(X_t|\mathcal{G}^{t-1}) = \alpha X_{t-1} \text{ a.s.}, \tag{2}$$

   for some unknown value α. Student A is unaware of (2) and estimates an ARMA(2,1) model. Her AR coefficient estimates are 0.29 and 0.11 and her MA coefficient estimate is 0.11. Take a guess at the approximate value of α with a **brief** explanation.

   (d) Professor O. estimates α to be 0.93. For $\alpha = 0.93$ $\{X_t\}$ is stationary, but Professor O. is hesitant to use the normal limiting distribution to make inferences about α. Should Professor O. indeed be concerned? Explain **briefly**.

   (e) Professor O. has another time series $\{Y_t\}$ covering the same time period and regresses Y_t on X_t. He then uses an Engle–Granger test.

   i. Explain **briefly** the workings of the Engle–Granger test.

   ii. Professor O. finds that the Engle–Granger test does not reject and becomes concerned once again. Why would he be?

2. Student B has a nice panel data set with $T > 2$ time periods and $n > 1000$ observations. Student B formulates a linear model explaining earnings y_{it} in terms of a number of demographic and job–related characteristics such as age, sex, education and position, which are gathered in a vector of regressors x_{it}.

   (a) As a first pass, student B runs OLS of y_{it} on x_{it} using all nT observations.

   i. Explain **briefly** what assumptions student B is implictly making for the OLS estimator to be consistent.

   ii. Suppose that the assumptions necessary for consistency of the OLS estimator are satisfied. Explain why the OLS estimator is then almost necessarily inefficient.

   (b) After meeting with his advisor, student B puts in a fixed effect r_i, which gives the model

$$y_{it} = x_{it}'\beta + r_i + u_{it}, \quad i = 1, \ldots, n, \ t = 1, \ldots, T. \tag{3}$$

   Student B then estimates β using the fixed effects estimator.

   i. Student B finds that he is no longer able to estimate all β–coefficients. Explain **briefly**.

   ii. Can you think of a way to recover the missing β–coefficients? Explain **briefly**.

   iii. Show that under certain conditions the fixed effects estimator is more efficient than the first differences estimator.

   (c) Student B then decides to add a lagged dependent variable as a regressor.

   i. Why is the fixed effects estimator no longer appropriate here? Explain **briefly**.

   ii. How can we still estimate the unknown coefficients? Argue **briefly** why your new estimator will be consistent.

[1] Known to some as 'information set'.

3. (a) Please answer the following true/false questions. No points without a correct and **brief** explanation.

   i. Choosing a larger bandwidth in nonparametric kernel regression estimation usually increases the estimator variance.

   ii. With higher order kernels one would generally choose a bandwidth which tends to zero faster than with a second order kernel.

   iii. Local polynomial estimators are better at estimating peaks and troughs of a regression function than are nonparametric kernel regression estimators.

   iv. The curse of dimensionality is related to how fast the nonparametric regression estimator bias goes to zero as the number of regressors increases.

   (b) Prove consistency of the nonparametric kernel density estimator.

   (c) Student C has an i.i.d. sample $\{(X_i, Y_i)\}$ and is interested in $m(x) = E(Y_1|X_1 = x)$. Student C has accidentally entered an incorrect, and very large, value for one of the Y_i's. Indicate in separate graphs what effect this mistake would have if student C (i) used a local nonparametric estimator and (ii) if (s)he used a global nonparametric estimator.

PART III

Mandatory Question

1. A representative household's lifetime utility is given by:

$$\max E \sum \beta^t \frac{\{c_t^{\alpha} l_t^{1-\alpha}\}^{1-\sigma} - 1}{1-\sigma}$$

subject to:

$$w_t h_t \geq c_t$$
$$h_t + l_t \leq 1$$

where c is consumption, l is leisure, or non-market time, β, α, σ are preference parameters, w is the wage rate, and the time endowment is normalized to be 1.

Given data on consumption, hours worked, and the real wage, describe how you would estimate the preference parameters $\beta, \alpha,$ and σ using GMM. Specifically, do the following:

Write down:

(i) the equations you would use to form the criterion funtion

(ii) how you might adjust the data for trends

(iii) the criterion function, including how you would form the weighting matrix

(iv) how you would optimize the criterion function.

(v) In what sense is GMM easier to use - and might give more reasonable results - than MLE?

Optional Question

2. Suppose that the measured inflation rate is a noisy indicator of the true inflation rate, which is a stationary process:

$$\pi_t = \pi_t^* + \varepsilon_t, E(\varepsilon) = 0, E(\varepsilon^2) = \sigma_\varepsilon^2$$
$$\pi_t^* = \rho_t m_t + v_t, E(\varepsilon) = 0, E(v^2) = \sigma_v^2$$
$$\rho_t = \bar{\rho} + \eta_t, E(\eta) = 0, E(\eta^2) = \sigma_\eta^2,$$

where m_t is the growth rate of the money supply.

Set this model up as a state space model, and describe how you would use the Kalman Filter to estimate the time-varying parameter ρ_t.