BOSTON COLLEGE
Department of Economics

Econometrics Field Comprehensive Exam

May 24, 2019

YOU MUST ANSWER PART A

PART A – ECON7772 –PROF. LEWBEL

AND THEN CHOOSE 2 PARTS OF THIS EXAM FROM THE FOLLOWING:

PART B – ECON8821 –PROF. XIAO

PART C – ECON8822 –PROF. HODERLEIN

PART D – ECON8823 –PROF. BAUM

PART E – ECON8825 –PROF. KHAN

Please use a separate bluebook for each part.

Please write your alias, part and question number(s) on each bluebook.

Please read the entire exam before writing anything.

Field Examination in Econometrics
May 2019 PART A

1. Suppose we have n i.n.i.d (independent, not identically distributed) observations of the random variable X_i, for $i = 1, ..., n$. Assume each X_i is normal with distribution $X_i \sim N(\mu, i)$. Consider the usual estimator $\widehat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Explain your answer to each of the following questions.

a. Is $\widehat{\mu}_n$ an unbiased estimator of μ?

b. Does $\widehat{\mu}_n$ converge in mean square to μ?

c. What is the distribution of $\widehat{\mu}_n$?

d. Does $\widehat{\mu}_n = O_p(1)$?

e. Does $\widehat{\mu}_n = O_p\left(n^{1/2}\right)$?

f. Does $\widehat{\mu}_n = o_p\left(n^{1/2}\right)$?

g. Is $\widehat{\mu}_n$ a consistent estimator of μ?

2. Consider the following system of equations:

$$\begin{aligned} Y_i &= aZ_i + bX_i + e_i \\ X_i &= cY_i + U_i \end{aligned}$$

Assume our data are independent, identically distributed observations of (Z_i, X_i, Y_i) for $i = 1, ..., n$. Assume $E\left[(Z_i, X_i, Y_i)\right] = 0$ and that all the elements of the matrix $E\left[(Z_i, X_i, Y_i)(Z_i, X_i, Y_i)'\right]$ are finite and nonzero. Assume the errors (e_i, U_i) are mean zero and uncorrelated with Z_i. The scalars a, b, and c, are unknown.

a. What are the reduced form equations for this model?

b. Either show that the coefficient b is not identified, or provide a consistent estimator for b.

c. Now assume that we know that $a = 1$, so now only b and c are unknown. Again either show that the coefficient b is not identified, or provide a consistent estimator for b.

3. Assume that W_i for $i = 1, ..., n$ are iid continuously distributed random variables. Let the cumulative distribution function of W_i be $F(W_i, \theta)$ and let the probability density function of W_i be $f(W_i, \theta)$. Assume we know the functions F and f, but we do not know the scalar parameter θ. We do not observe W_i, but we do observe X_i for $i = 1, ..., n$, where $X_i = W_i I(W_i \geq 0)$. So X_i is zero if W_i is negative, otherwise X_i equals W_i. What is the log likelihood function for estimating θ, using the data X_1,...,X_n? What is the score function?

PART B

Time Series Econometrics, May, 2019

Question 1. (i) Given a MA(2) process

$$X_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}, \text{ where } \varepsilon_t \sim WN(0, \sigma^2),$$

and information upto time n, find the optimal linear forecast of X_{n+h}, $h > 0$, forecast errors, and forecast error variance. (ii) Given a MA(∞) process

$$X_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}, \text{ where } \theta_0 = 1,\ \varepsilon_t \sim WN(0, \sigma^2),$$

and information upto time n, find the optimal linear forecast of X_{n+h}, $h > 0$, forecast errors, and forecast error variance..

Question 2. Consider a linear Regression with correlated errors:

$$y_t = \beta' x_t + u_t,$$

where (i) $\{x_t\}$ and $\{u_t\}$ are independent with each other (ii) (x_t, u_t) is α-mixing with size $-r/r-1$ ($r > 1$), (iii) $\sup_{it} E|x_{it}u_t|^{r+\delta} < \infty$, (iv) $\sup_{it} Ex_{it}^{2(r+\delta)} < \infty$, $M = \frac{1}{n}\sum_1^n E(x_t x_t')$ is uniformly positive definite (i.e., $M_n \geq \varepsilon I > 0$ for all large n).

Prove that the OLS estimator for β is consistent and asymptotic normal.

Question 3. Consider the AR process

$$y_t = \alpha y_{t-1} + u_t,\ \alpha = 1$$

where u_t are stationary but weakly dependent process, say, a mixing process whose partial sums satisfies a functional central limiting theorem, and consider the OLS estimation of the AR coefficient

$$\widehat{\alpha} = \frac{\sum y_{t-1} y_t}{\sum y_{t-1}^2}$$

what's the limiting distribution of $\widehat{\alpha}$? (ii) How to construct an unit root test based on the OLS estimator?

PART C

EC 8822

Stefan Hoderlein

May, 24 2019

Instructions: this part is an one hour, closed book exam. Attempt to answer formally and avoid lengthy verbal parts. There are 40 credits in total, and the individual credits for the questions are displayed as guidance. Try to answer all questions. However, the best grade is usually awarded with less than full credit.

1. State the Kernel based estimator for the density f_X of a d-dim random vector X. Define all elements formally, including the Kernel ("a Kernel is a function K such that...."). Explain how the estimator differs from a Kernel based estimator, if X had only one dimension. (6 credits)

2. Based on your answers in question 1, how would you estimate the mean regression of a random variable Y on a d-random vector X denoted $m(x)$? Hint: use the fact that in the population

$$m(x) = E[Y|X = x] = \int y f_{Y|X}(y; x) dy,$$

and derive a sample counterparts estimator. (4 credits)

3. Can you run a nonparametric regression if Y is binary? If yes, why, if not, why not? (2 credits)

4. What type of restrictive assumption necessary for (parametric) maximum likelihood based estimation of the binary choice model with d dimensional X is not required for nonparametric estimation of the binary choice model? What is the drawback of a general kernel based estimator when compared to parametric maximum likelihood estimation (hint: focus on the rates of convergence of the two estimators)? (3 credits)

5. In which sense is a semiparametric binary choice model a good compromise between the two paradigms? (2 credits)

6. Consider the local average treatment effects (LATE) when there is a binary treatment D and a binary instrument Z, such that Z is independent of all unobservables in the system. Define this quantity, and explain how it can be identified (just the identification result, no proof). Why is this quantity instrument dependent, and what are associated disadvantages? (5 credits)

7. The following question applies to the linear random coefficients model

$$Y = B_0 + X'B_1,$$

where Y is an observed random scalar, X is an observed random K-vector, $B = (B_0, B_1)'$ is an unobserved random $K+1$ vector, and X is fully independent of B. The goal is to obtain an estimator for the density f_B. To this end, define the conditional characteristic function of $Y|X = x$ as

$$\varphi_{Y|X}(t; x) = E\left[\exp(itY)|X = x\right],$$

where $t, x \in R^{K+1}$, and $i = \sqrt{-1}$.

Show that there is a one-to-one relation between $\varphi_{Y|X}(t; x)$ and $\varphi_B(s)$ for any value of $(t, x, s) \in R^{2K+2}$ (4 credits)

8. What is a regression (random) tree? Explain its functioning at every single node. (3 credits)

9. What is pruning and why is it used? Describe an algorithm that estimates a tree. (3 credits)

10. Define the Lasso and the Ridge estimator by stating the optimization problem on which both estimators are based. Define all quantities (4 credits).

11. What is the difference between the two approaches? Focus on variable selection, and explain how this feature relates to the difference in the optimization problem you described in question 10. Hint: use a detailed graphical explanation, **clearly defining all objects in the graph** (4 credits).

Good Luck, Guys!!!

PART D

ECON 8823: Applied Econometrics Comprehensive, Spring 2019

Answer all questions. They are equally weighted.

1. Write an essay on the choice of instruments in an instrumental variables estimation problem, discussing exact vs. overidentification and the 'weak instruments' problem.

2. Write an essay discussing the concept of long memory, or fractional integration of a time series, describing how it could be detected and estimated.

3. Write an essay discussing the econometric methods that could be applied to model count data, such as the number of firearm homicides per month and city.

PART E

Econ 8825 May 21, 2019

Spring 2019

1. (30) This question pertains to the multinomial choice model studied in class.

   Assume the dependent variable takes one of $J+1$ mutually exclusive and collectively exhaustive alternatives (numbered from $j=0$ to $j=J$).

   Recall the model was based on utility maximization. Specifically, for individual i, alternative j is assumed to have an unobservable indirect utility y_{ij}^* and the alternative with the highest indirect utility is assumed chosen.

   Thus the observed variable y_{ij} has the form

   $$y_{ij} = I[y_{ij}^* > y_{ik}^* \text{ for } k = 0, ..., J]$$

   An assumption of joint continuity of the indirect utilities rules out ties (with probability one);

   In the model for this question, the indirect utilities are further restricted to have the linear form

   $$y_{ij}^* = x_{ij}'\beta_0 + \epsilon_{ij}$$

   for $j = 0, ..., J$, where the $J+1$ dimensional vector ϵ_i (with elements ϵ_{ij}) of unobserved error terms is assumed to be jointly continuously distributed and independent of the $(J+1) \times k$-dimensional matrix of regressors X_i (whose j^{th} row is x_{ij}).

   The purpose of this question is to focus on identification and estimation of β_0, from a random sample of observations of the vector (y_i, x_i), where $y_i \equiv (y_{i0}, ...y_{iJ}), x_i \equiv (x_{i0}, ...x_{iJ})$.

   We will first take a parametric approach.

   (a) Assume the vector $\epsilon_{ij}, j = 0, 1, ...J$ is multivariate normal and distributed independently of the regressors x_{ij}. List all scale and location normalizations needed to identify β_0.

   (b) Write down the likelihood function for this model.

   (c) Propose an algorithm to construct a simulated MLE (SMLE). What motivates an SMLE over an MLE?

2. (30) Consider a variation of the standard *censored* selection model, where we get to observe both hours worked and wages of individuals. Hours worked are censored at 0 and we only observe wages for those who are working, and hence their hours worked is positive.

$$
\begin{aligned}
y_{1i} &= \max(w_i'\delta_0 + \eta_i, 0) \\
d_i &= I[y_{1i} > 0] \\
y_i &= d_i(x_i'\beta_0 + \epsilon_i)
\end{aligned}
$$

Observed variables: $(y_{1i}, d_i, w_i, y_i, x_i)$ and we assume $(\eta_i, \epsilon_i) \sim N(\mathbf{0}, \Sigma)$, where the (1,1) element of Σ is σ_1^2, the other diagonal element is σ_2^2, and the off diagonal terms are γ_0. We wish to estimate $(\delta_0, \beta_0, \gamma_0, \sigma_1^2, \sigma_2^2)$.

(a) Write the likelihood function for this model.

(b) Propose a NLLS estimator for the parameters in this model.

(c) Propose a computationally friendly Heckman type two-step estimator where the first step involves a convex optimization procedure and the second step is closed form.

3. (15) Frechet(1951) was interested in knowing whether knowing the marginal distributions of continuous random variables X and Y tells us anything about their joint distribution.

He showed that, given knowledge of the distribution functions $F(\cdot)$ and $G(\cdot)$ of X and Y, respectively, their joint distribution $K(\cdot,\cdot)$is such that for all a, b:

$$\max(F(a) + G(b) - 1, 0) \leq K(a, b) \leq \min(F(a), G(b))$$

Prove this result.