Department of Economics

UCLA

Econometrics Field Exam

Spring 2012

This exam consists of 3 questions in Part I, 3 questions in Part II, and 5 questions in Part III. You are required to answer all the questions. Good luck!

Part I

Question 1

Consider the following model

$$Y = \sum_{k=1}^{2} s_k \left(X_k, \varepsilon_k \right)$$

where $(Y, X_1, X_2) \in R^{1+L_1+L_2}$ is observable, $(\varepsilon_1, \varepsilon_2)$ is unobservable, and each of the unknown functions s_k $(k = 1, 2)$ is continuously differentiable in (X_k, ε_k) and strictly increasing in ε_k. Assume that (X_1, X_2) is distributed independently of $(\varepsilon_1, \varepsilon_2)$, that the support of (X_1, X_2) is $R^{L_1+L_2}$ and that the support of $(\varepsilon_1, \varepsilon_2)$ is R^2.

a. Determine whether the functions s_1, s_2 and the distribution of $(\varepsilon_1, \varepsilon_2)$ are identified. If your answer is YES, prove it. If your answer is NO, impose sufficient additional conditions under which the functions and the distribution are identified, and show identification under those conditions.

b. Given i.i.d. data $\{(Y_1^i, X_1^i, X_2^i)\}_{i=1}^N$ generated from the above model, propose consistent estimators for the functions s_1, s_2 and the distribution of $(\varepsilon_1, \varepsilon_2)$. Describe the main steps you would follow to show that your proposed estimators are consistent.

Question 2

Suppose that the model is

$$\begin{aligned} Y_1 &= m\left(Y_2\right) + \eta_1 \\ Y_2 &= \gamma Y_1 + \beta' Z + \eta_2 \end{aligned}$$

where the unknown function m is continuous differentiable, the values of the parameters γ and β are unknown, and the distribution of (η_1, η_2) is also unknown. Assume that the vector (Y_1, Y_2, Z) is observable and has support R^3, that the vector (η_1, η_2) is unobservable and has support R^2, and that Z and (η_1, η_2) are distributed independently of each other.

Determine whether the derivative of m is identified. If your answer is YES prove it. If your answer is NO, provide a set of additional conditions under which the derivative is identified, and prove that it is identified under those conditions.

Question 3

Consider the model

$$Y_1 = b(X) + \eta$$

where $(Y, X) \in R^2$ is observable, η is unobservable, X and η are independently distributed, $E(\eta) = 0$, the support of X is the set of points $\{x_1, ..., x_M\}$, and the unknown function $b : R \to R$ is concave.

a. Determine whether the function b and the distribution of η are identified. If your answer is YES prove it. If your answer is NO, explain your answer, determine which values of b and of the distribution of η are point identified, and determine sharp bounds for the values of b that are not point identified.

b. Propose a consistent estimator for the identified values of b and state the main steps you would follow to prove the consistency.

Part II

Question II.1

Consider a model of social interaction between two roommates:

$$\begin{aligned} y_{i1} &= \alpha_i + x_{i1}\beta + y_{i2}\delta + \varepsilon_{i1} \\ y_{i2} &= \alpha_i + x_{i2}\beta + y_{i1}\delta + \varepsilon_{i2} \quad i = 1, \ldots, n \end{aligned}$$

where y_{i1} denotes the first student's GPA in the ith dormitory room, x_{i1} denotes his/her own observed academic background, ε_{i1} denotes his own unobserved characteristic. The α_i denotes the unobserved room specific factor (such as humidity or temperature). We will assume that the room and roommates are randomly assigned, which implies that (x_{i1}, x_{i2}) is independent of α_i, and x_{i1} and x_{i2} are independent of each other with common distribution. Finally, we will assume that α_i and $(\varepsilon_{i1}, \varepsilon_{i2})$ are independent of each other, and (x_{i1}, x_{i2}) and $(\varepsilon_{i1}, \varepsilon_{i2})$ are independent of each other. Assume that there are lots of rooms in the dorm, and we will assume the asymptotics where $n \to \infty$. Prove that δ is identified and present a consistent estimator. (You can assume that $|\delta| < 1$.) Does your estimator remain consistent if you drop the assumption that (x_{i1}, x_{i2}) is independent of α_i?

Question II.2

Consider a panel logit model with fixed effects:

$$\begin{aligned} \Pr(y_{i1} = s, y_{i2} = t | x_{i1}, x_{i2}, \alpha_i) = &\left[\frac{\exp(\alpha_i + x_{i1}\beta)}{1 + \exp(\alpha_i + x_{i1}\beta)}\right]^s \left[\frac{1}{1 + \exp(\alpha_i + x_{i1}\beta)}\right]^{1-s} \\ &\times \left[\frac{\exp(\alpha_i + x_{i2}\beta)}{1 + \exp(\alpha_i + x_{i2}\beta)}\right]^t \left[\frac{1}{1 + \exp(\alpha_i + x_{i2}\beta)}\right]^{1-t} \end{aligned}$$

for $s, t = 0, 1$. Assume that $(y_{i1}, y_{i2}, x_{i1}, x_{i2}, \alpha_i)$ $i = 1, \ldots, n$ are i.i.d. Assume that you only observe $(y_{i1}, y_{i2}, x_{i1}, x_{i2})$. Propose a consistent estimator of β.

Question II.3

Consider a linear regression model

$$\begin{aligned} y_{gi} &= \beta\pi_g + \varepsilon_{gi} \\ x_{gi} &= \pi_g + v_{gi} \quad g = 1, \ldots, G; i = 1, \ldots, K \end{aligned}$$

Adopt an asmyptotic approximation that $G \to \infty$ while K is fixed. Discuss the asymptotic property of the natural estimator that regresses y_{gi} on $\overline{x}_g = \frac{1}{K}\sum_{i=1}^{K} x_{gi}$. If problematic, propose an estimator that overcomes the problem that you identified.

Part III

Question III.1

Suppose the following assumptions hold:

a. $Y = D \cdot Y^1 + (1 - D) \cdot Y^0$, where $D \in \{0, 1\}$ and $Y \in [0, 1]$,
b. $D = Z \cdot D^1 + (1 - Z) \cdot D^0$, where $Z \in \{0, 1\}$,
c. $Z \perp (Y^0, Y^1, D^0, D^1)$,
d. $D^1 \geq D^0$,
e. $Y^1 \geq Y^0$.

Suppose we are interested in the effect of treatment on the difference between the 90^{th} and the 10^{th} percentile of the outcome distribution,

$$\nu(Y) = F_Y^{-1}(.9) - F_Y^{-1}(.1).$$

What is the identified set for the effect of treatment on ν, $\nu(Y^1) - \nu(Y^0)$?

Question III.2

A measure of inequality that has received increasing public attention recently is the share of income going to the top 1% of income earners. Denoting by F be the c.d.f. of the (continuous) income distribution, this parameter can be written as

$$\nu(F) = \frac{\int_{F^{-1}(.99)}^{\infty} y \, dF(y)}{\int_0^{\infty} y \, dF(y)}.$$

Calculate the influence function of ν.

Queston III.3

A researcher is interested in the effect of urban segregation on crime rates. In particular, she hypothesizes that crime is affected by neighborhood-level social interactions. Building on (Graham, Imbens, and Ridder 2008), she considers a model in which the equilibrium crime rate Y in a neighborhood is given by

$$Y = m(X, U),$$

where X is the poverty rate in the neighborhood and U are further unobserved determinants of crime. She is interested in the effect on aggregate crime $E[Y]$ of an intervention that increases segregation, while holding average poverty constant. In particular, X is replaced by $X + \alpha \cdot (X - E[X])$ for all neighborhoods, while holding U constant. α is a positive constant.

In order to identify this effect, she uses an instrument Z. X is related to Z through

$$X = h(Z, V),$$

where it is assumed that h is strictly monotonic in the one-dimensional V, and $Z \perp (U, V)$.

Characterize the identified set for the effect of the policy intervention on aggregate crime. You can assume that Y is bounded by $[0, \overline{Y}]$. Discuss the role of the joint support of (X, Z) for identification. Drawing a figure might help.

Question III.4

Suppose we observe data on income Y and education X of women (P^1) and men (P^0). We are interested in the counterfactual income distribution

$$\tilde{P}(Y) = \int P^1(Y|X)P^0(X)dX.$$

As discussed in class, how can we rewrite $\tilde{P}$ as a reweighted baseline distribution, using weights $w(X)$?

Now consider the following three scenarios, where in each case we are given the probability densities f^1 and f^0 of education for women and men:

1. $f^0(x) = \mathbf{1}(x \in [0, 1])$,
   $f^1(x) = \frac{1}{2} \cdot \mathbf{1}(x \in [0, 2])$
2. $f^0(x) = 2 \cdot \mathbf{1}(x \in [0, 1]) \cdot x$,
   $f^1(x) = 3 \cdot \mathbf{1}(x \in [0, 1]) \cdot x^2$
3. $f^0(x) = \mathbf{1}(x \in [0, 1])$,
   $f^1(x) = 2 \cdot \mathbf{1}(x \in [.5, 1])$

For each of these cases, discuss identifiability of the counterfactual distribution $\tilde{P}$, and the behavior of the weight function w. How do you think does the behavior of this weight function relate to potential difficulties in estimation?

Question III.5

A policy maker has the objective of maximizing revenue Y from the top tax-bracket, through choice of the top tax-rate X. She assumes a quadratic model for the relationship between the tax rate and revenues,

$$Y = \beta_0 + \beta_1 X + \frac{\beta_2}{2} X^2 + U,$$

where $U \sim^{i.i.d.} N(0, \sigma^2)$. She knows β_2 and has a joint normal prior for the other parameters,

$$(\beta_0, \beta_1) \sim N(\mu, \sigma).$$

Suppose furthermore that she observes n draws from the joint distribution of X and Y, where it is assumed that $X \perp U$.

Under these assumptions,

a. What is the posterior expectation of β_0 and β_1?

b. What is the tax rate x^* maximizing expected revenues, given the data?

c. What are expected revenues at the optimal tax rate x^*?