**Department of Economics**
**UCLA**

**Econometrics Field Exam**
**Fall, 2010**

# Second Year Field Examination in Econometrics
# September 10, 2010

There is a total of seven questions, three in Part I, two in Part II, and two in Part III. Please answer six out of the seven available questions.

Please provide as detailed information as possible. please keep in mind that the answers need not be long for them to be precise.

## Part I:

### Question 1:

Assume the following:

1. There are four potential values $Y_0, Y_1$ and $D(0), D(1)$. We also observe an instrument $Z$, which is binary.
2. $Y_1$ denotes the potential outcome under treatment, and $Y_0$ denotes the potential outcome under control.
3. $D(1)$ denotes the potential treatment when $Z = 1$, and $D(0)$ denotes the potential treatment when $Z = 0$.
4. We assume that $(Y_0, Y_1, D(0), D(1))$ is independent of $Z$.
5. We also assume that $D(0) \leq D(1)$.

Prove the following equality:

$$E[Y_1 - Y_0 | D(1) - D(0) = 1] = \frac{E[Y | Z = 1] - E[Y | Z_i = 0]}{E[D | Z = 1] - E[D | Z = 0]}$$

### Question 2:

Consider the linear regression model

$$y_i = x_i \beta + \varepsilon_i$$

where it is known that $E[\varepsilon_i | z_i] = 0$. (For simplicity, we will assume that every random variable is a scalar.) This implies that

$$E[h(z_i)(y_i - x_i \beta)] = 0$$

where $h(z_i) = E[x_i | z_i]$. Consider implementing this moment restriction for estimating $\beta$, i.e., consider $\widehat{\beta}$ that solves

$$\frac{1}{n} \sum_{i=1}^{n} \widehat{h}(z_i) \left( y_i - x_i \widehat{\beta} \right) = 0$$

where $\widehat{h}(z_i)$ is a nonparametric estimator of $h(z_i) = E[x_i|\, z_i]$. Using the Newey formula, find the influence function of $\sqrt{n}\left(\widehat{\beta} - \beta\right)$, and characterize the asymptotic variance.

**Question 3:**

Consider the standard panel model

$$y_{it} = \alpha_i + \varepsilon_{it}, \quad i = 1, \ldots, n; t = 1, \ldots, T$$

where $\varepsilon_{it}$ is iid $N(0, \theta_0)$. Derive the MLE for $\theta$, and find its probability limit as $n \to \infty$ while $T$ is fixed. Let $\theta_T$ denote such a probability limit. Show that $\theta_T$ allows the expansion of the form

$$\theta_T = \theta_0 + \frac{B}{T} + o\left(\frac{1}{T}\right)$$

What is $B$?

## Part II:

### Question 1:

Consider the model

$$Y = m(X, \varepsilon)$$

where $Y \in R$ and $X \in R^K$ are observable, $\varepsilon \in R$ is unobservable, and $m$ is differentiable in $x$ and continuous in $\varepsilon$. For any random vector $W$, let $f_W(w)$ denote the density of $W$ at $w$ and let $f_{W_1|W2=w_2}(w_1)$ denote the conditional density of $W_1$ at $W_1 = w_1$ given that $W_2 = w_2$.

Suppose that you are interested in

$$\beta(x) = \int \frac{\partial m(x, \varepsilon)}{\partial x} \; f_{\varepsilon|X=x}(x) \; dx$$

1. Provide an estimator for $\beta(x)$ when $\varepsilon$ is distributed independently of $X$. What can you say about the asymptotic properties of the estimator that you provided? Explain.

2. Provide an estimator for $\beta(x)$ when $\varepsilon$ is not distributed independent of $X$, but for some observable $Z$, it is known that for all $x, \varepsilon, z$,

$$\frac{f_{\varepsilon,X,Z}(\varepsilon, x, z)}{f_Z(z)} = \frac{f_{\varepsilon,Z}(\varepsilon, z)}{f_Z(z)} \frac{f_{X,Z}(x, z)}{f_Z(z)}$$

   What can you say about the asymptotic properties of the estimator that you provided? Explain.

3. Provide an estimator for $\beta(x)$ when (i) $\varepsilon$ is not distributed independently of $X$, but (ii) for an unknown function $s$, which is strictly increasing in an unobserved $\eta$, and for an observed $Z$ that is distributed independently of $(\varepsilon, \eta)$,

$$X = s(Z, \eta).$$

   What can you say about the asymptotic properties of the estimator that you provided? Explain.

### Question 2:

Suppose that the model that determines the distribution of $(Y_1, Y_2)$ given $(X_1, X_2)$ is

$$\begin{aligned} Y_1 &= m_1(Y_2, \varepsilon_1 + X_1) \\ Y_2 &= m_2(Y_1, \varepsilon_2 + X_2) \end{aligned}$$

where $m_1$ and $m_2$ are continuous functions, both strictly increasing in their last coordinates, and $(\varepsilon_1, \varepsilon_2)$ is an unobserved random vector of unknown everywhere strictly positive density. Assume that $(X_1, X_2)$ is distributed independently of $(\varepsilon_1, \varepsilon_2)$ and the distribution of $(Y_1, Y_2, X_1, X_2)$ is observed. What can you say about the identification of the derivatives of $m_1$ and $m_2$? Provide details.

## Part III:

### 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 the amount of time invested in acquiring human capital, say $h_t^h$. Let also $h_t^l$ denotes the number of hours worked. Note that

$$l_t + h_t^h + h_t^l = L,$$

where $L$ is a fixed number of hours.

Each period utility is given by

$$u_t(c_t, l_t) = u(c_t, l_t) = K_0 c_t^{\gamma_1} l_t^{\gamma_2},$$

In each period the individual obtains a fixed amount (known in advance) of unearned income, denoted $a_t$. If the individual works then he/she is paid an hourly wage $w_t$, which is a function of the level of human capital $H_t$, that is,

$$\log w_t = \alpha_0 + \alpha_1 H_t + \varepsilon_t,$$

where $\varepsilon_t$ is an idiosyncractic shock and $H_t$ is a measured level of human capital that is evolved according to

$$H_t = H\left(H_{t-1}, I_t, h_t^h\right),$$

where $I_t$ is the monetary investment in human capital.

The individual starts with an endowment of $B_0$, 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. Make sure to provide the value function for *all* possible state of nature and choices.
4. Provide the full list of parameters that need to be estimated. Let the vector containing the full set of the model's parameters be denoted by $\theta_0$.
5. Provide details about the data that one would need in order to be able to estimate $\theta_0$.
6. Provide detailed information about a method by which you propose to estimate the parameter vector $\theta_0$. Keep in mind that there might be more than one method for estimating $\theta_0$. You should answer this question providing very detailed information, assuming you are giving instructions to a professional programmer who knows nothing about economics.
7. Suppose now that you want to consider a policy experiment in which the government will subsidize investment in human capital by offering reimbursement of 50% on the amount invested (i.e., $I_t$). How would you go about examining the impact of such policy?

**Question 2—Generalized Method of Moments:**

Consider the model given by

$$Y_i = g(X_i; \theta_0) + \varepsilon_i,$$

where the data given to you is $W_i = (Y_i, X_i, Z_i)$, $Z_i$ is $K_z \times 1$ vector of exogenous variables, and $i = 1, ...N$. Define the moment function

$$\varphi(Y_i, X_i, \theta) = (Y_i - g(X_i; \theta_0)) \, h(Z_i). \tag{2.1}$$

1. Under what conditions would the parameter vector $\theta_0$ in (2.1) be identified.
2. Suggest an efficient Generalized Method of Moments (GMM) estimator for $\theta_0$.
3. Suppose now that the function $g(X; \theta)$ cannot be analytically computed, but it can be approximated by a simulated version, say $\widehat{g}^R(X; \theta)$, where $R$ denotes the number of draws used from the relevant distribution to compute $\widehat{g}^R(X; \theta)$. Describe in details how to obtain the simulated GMM estimate for $\theta_0$.
4. Suppose now that you have estimated $\theta_0$, based on (3) and you want construct a symmetric confidence interval for $\theta_{02}$, the second element in $\theta_0$, using the bootstrap method. Describe how obtain such confidence interval. Be specific about the way the bootstrap samples are obtained, as well as the procedure followed for each bootstrap sample.