Second Year Field Examination in Econometrics
June 22, 2011

There is a total of six questions, two in each part of the exam. Please answer four out of the six available questions and at least one question from each of the first two parts, i.e., Part I and Part II.

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

Suppose that you are interested in estimating the functions \( m \) and \( s \) and the distribution of \((\varepsilon_1, \varepsilon_2)\) in the following model:

\[
\begin{align*}
y_1^* &= z + m(y_2, \varepsilon_1), \\
y_2 &= s(x, \varepsilon_2), \\
y_1 &= \begin{cases} 
1 & \text{if } y_1^* \geq 0 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

where \( y_1, y_2, x, \) and \( z \) are observable random variables, \((\varepsilon_1, \varepsilon_2)\) and \( y_1^* \) are unobservable, the functions \( m \) and \( s \) are differentiable, and the distribution of \((\varepsilon_1, \varepsilon_2)\) is also differentiable.

Assume that (i) \( z, x, \) and \((\varepsilon_1, \varepsilon_2)\) are mutually independent, (ii) \( m \) is strictly increasing in \( \varepsilon_1 \), (iii) \( s \) is strictly increasing in \( \varepsilon_2 \), (iv) the support of \( z \) is \( R \), (v) the support of \( x \) is \( R \), and (vi) the support of \((\varepsilon_1, \varepsilon_2)\) is \( R^2 \).

1. Imposing additional conditions when necessary, obtain identification results for the function \( s \), the function \( m \), and the distribution of \((\varepsilon_1, \varepsilon_2)\).

2. Justify the additional conditions that you used in (1).

3. Describe a nonparametric method to estimate the functions \( s \) and \( m \), under the conditions that you imposed in (1) and additional ones if necessary.

4. Describe the asymptotic properties of the estimator that you provided in (3).

Question 2

Consider the following model

\[
\begin{align*}
y_1 &= m^1(y_2, \beta_1^1 x_1 + \varepsilon_1), \\
y_2 &= m^2(y_1, \beta_2^2 x_2 + \varepsilon_2),
\end{align*}
\]

where the functions \( m^1 \) and \( m^2 \) are differentiable unknown functions, each strictly increasing in its last coordinate, \((x_1, x_2)\) is distributed independently of \((\varepsilon_1, \varepsilon_2)\), and the support of \((\varepsilon_1, \varepsilon_2)\) is \( R^2 \).
1. Imposing additional conditions as necessary, provide conditions under which \( \beta_1 \) and \( \beta_2 \) are identified.

2. Imposing even more additional conditions as necessary, provide conditions under which the derivatives of \( m^1 \) and \( m^2 \) are identified.

3. Provide estimators for \( \beta_1 \) and \( \beta_2 \) and discuss their asymptotic properties.

4. Provide estimators for the derivatives of \( m^1 \) and \( m^2 \) and discuss their asymptotic properties, under additional assumptions if necessary.

**Part II:**

**Question 1**

Consider the model

\[ y = x \beta + \varepsilon, \]

where we know \( E[\varepsilon | z] = 0 \). We use the following two-step estimator of \( \beta \):

1. Estimate \( E[x | z] \) by nonparametric regression \( \hat{E}[x | z] \).
2. Compute the OLS of \( y \) on \( x \) and \( \hat{\beta} = x - \hat{E}[x | z] \).
3. The estimator \( \hat{\beta} \) of \( \beta \) is the OLS coefficient of \( x \) in the second step. Assuming that \( \hat{\beta} \) is consistent, find the asymptotic variance of \( \hat{\beta} \) using Newey’s (1994) formula.

**Question 2**

Adopt the following notations:

- For each individual, there are four potential values \( Y_0, Y_1 \) and \( D(0), D(1) \). We also observe an instrument \( Z \), which takes two values, \( \underline{z} \) and \( \overline{z} \).
- \( Y_1 \) denotes the potential outcome under treatment, and \( Y_0 \) denotes the potential outcome under control.
- \( Z \) takes two possible values, \( \underline{z} \) and \( \overline{z} \).
- \( D(\overline{z}) \) denotes the potential treatment when \( Z = \overline{z} \), and \( D(\underline{z}) \) denotes the potential treatment when \( Z = \underline{z} \).
- We observe \( Z, D = D(Z) \), and \( D(Z) Y_1 + (1 - D(Z)) Y_0 \)
- We assume that \( (Y_0, Y_1, D(\underline{z}), D(\overline{z})) \) is independent of \( Z \)
- We assume that \( D(\underline{z}) \leq D(\overline{z}) \).

1. Prove that

\[
E[Y_1 - Y_0 | D(\overline{z}) - D(\underline{z}) = 1] = \frac{E[Y | Z = \overline{z}] - E[Y | Z = \underline{z}]}{E[D | Z = \overline{z}] - E[D | Z = \underline{z}]}.
\]
2. Now, suppose that

\[
\lim_{\tau \to \infty} \Pr [D (\tau) = 1 | \tau] = 1, \\
\lim_{z \to \infty} \Pr [D (z) = 0 | z] = 1.
\]

Prove that

\[
\begin{align*}
E [Y_1 - Y_0] &= \lim_{\tau \to \infty, z \to \infty} \frac{E [Y | Z = \tau] - E [Y | Z = z]}{E [D | Z = \tau] - E [D | Z = z]}, \\
&= \lim_{\tau \to \infty, z \to \infty} \frac{E [Y_1 - Y_0]}{E [D | Z = \tau] - E [D | Z = z]},
\end{align*}
\]

Part III:

Question 1—Dynamic Programming:

Consider a finite horizon discrete dynamic programming model in which an individual makes decisions about three variables: consumption \(c_t\), leisure \(l_t\), and investment in health \(I^h_t\). An individual may decide not to work, in which case he/she will consume the full amount of available leisure.

Each period utility is given by

\[
u_t (c_t, l_t, h_t) = u (c_t, l_t, h_t) = K_0 c_t^{\gamma_1} l_t^{\gamma_2} h_t^{\gamma_3}.
\]

Each period an individual gets an amount of unearned income \(a_t\). If in addition, the individual works, then he/she is paid an hourly wage \(w_t\), which is a function of the number of years of education denoted \(e\), and experience, denoted \(x\). Specifically, we have

\[
\log w_t = \alpha_0 + \alpha_1 e_t + \alpha_2 e_t^2 + \alpha_3 x_t + \alpha_4 x_t^2 + \alpha_5 x_t e_t + \varepsilon_t,
\]

where \(\varepsilon_t\) is an idiosyncratic shock, uncorrelated with either education or experience. The current level of health is given by

\[
h_t = \exp \left\{ \delta_0 + \delta_1 h_{t-1} + \delta_2 I^h_t + \nu_t \right\}.
\]

The individual starts with an endowment of \(B_0\), and earns 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 necessary information has been omitted, please make assumptions about that information.

1. Define the state vector, say \(z_t\).

2. Specify all the necessary budget constraints, regarding money and time.

3. Specify the Bellman representation of the value function. Make sure to provide the value function for all possible states of nature and choices.

4. Provide the full list of parameters that need to be estimated. Let the true parameter vector 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 the method by which you propose to estimate the parameter vector \( \theta_0 \). You should answer this question providing enough detail so that a professional programmer who knows nothing about economics would be able to program the procedure. (Keep in mind that there are several alternative ways for obtaining an estimate for \( \theta_0 \ ).

7. What is the asymptotic distribution for the parameter vector estimate, say \( \hat{\theta}_n \), obtained following the procedure suggested in part (6)? Provide brief justifications for all your claims.

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

Consider the two moment functions given by \( \varphi_1(W_i; \theta_1) \) and \( \varphi_2(W_i; \theta_1, \theta_2) \). Suppose that, when evaluated at the true population parameter vectors, \( \theta_{01} \) and \( \theta_{02} \), respectively, we have

\[
E[\varphi_1(W_i; \theta_{01})] = 0 \quad \text{and} \quad E[\varphi_2(W_i; \theta_{01}, \theta_{02})] = 0,
\]

where \( W_i, i = 1, \ldots, n \), represent cross-section independent and identically distributed (i.i.d.) data.

Let \( \varphi_1(\cdot) \) and \( \varphi_2(\cdot) \) be \( M_1 \times 1 \) and \( M_2 \times 1 \) vector-valued functions, respectively, and correspondingly let \( \theta_{01} \) and \( \theta_{02} \) be \( d_1 \) and \( d_2 \) vectors of parameters, with \( M_1 > d_1 \) and \( M_2 = d_2 \).

Assume that you have already obtained an estimator for \( \theta_{01} \), say \( \hat{\theta}_{n1} \), and this estimator is consistent and asymptotically normal, i.e.,

\[
\hat{\theta}_{n1} \overset{p}{\rightarrow} \theta_{01}, \\
\sqrt{n} \left( \hat{\theta}_{n1} - \theta_{01} \right) \overset{D}{\rightarrow} N(0, V_1),
\]

as \( n \to \infty \).

1. Suppose that it has been determined that \( \theta_{01} \) has to be computed using the simulated method of moments (SMM) method. What would be the asymptotic distribution of the estimator for \( \theta_{01} \), say \( \hat{\theta}_{n1} \)? Justify each and every step in your answer.

2. Suggest an efficient Generalized Method of Moments (GMM) estimator for \( \theta_{02} \), using a plug-in estimator for \( \theta_{01} \), namely \( \hat{\theta}_{n1} \).

3. Show that the estimator suggested in (2) is consistent and has an asymptotic normal distribution. In doing so explicitly state all the assumptions you make. Provide the asymptotic properties of the estimator \( \hat{\theta}_{n2} \) for \( \theta_{02} \). Justify each and every step in your answer.

4. Suppose now that one would like to develop a test for the null hypothesis \( H_0: h_1(\theta_{01}) - h_2(\theta_{02}) = 0 \). Suggest a test and give details regarding the exact procedure you would follow. Be sure to give enough details so one can apply the procedure to a specific data.