UC Berkeley **Spring Semester 2004**

Economics 241A **Guido Imbens**

PROBLEM SET II, part 1,

Due: Wednesday March 3rd

1. (a) Consider the following two moment restrictions, for scalar random variables X and Y.

$$E[\psi(X, Y, \beta)] = 0,$$

where

$$\psi(x, y, \beta) = \begin{pmatrix} x(y - x\beta) \\ x - 3 \end{pmatrix}.$$

Suppose that in fact $Y|X$ is normal with mean βX and unit variance. What is the efficiency gain from using both moment restrictions relative to using only a single one?

(b) Answer the same question when the moment function is

$$\psi(x, y, \beta) = \begin{pmatrix} x(y - x\beta) \\ y - 3 \end{pmatrix}.$$

2. Use the data in the ascii file PROBII.DAT. In the census data the average of log earnings is 2.0647, the average of years of education 13.97, the average of log earnings squared is 4.5258, and the average of log earnings times education is 29.099. Because the census sample is so much larger we will ignore the sampling error in these estimates. Consider the moment restrictions

$$E[\psi(\ln(\text{earn}), \text{educ}, \text{exper}, \text{iq}, \beta^*)] = 0,$$

where

$$\psi(\ln(\text{earn}), \text{educ}, \text{exper}, \text{iq}, \beta)$$

$$
= \begin{pmatrix}
\log(\text{earn}) - \beta_0 - \beta_1 \cdot \text{educ} - \beta_2 \cdot \text{exper} - \beta_3 \cdot \text{exper}^2 - \beta_4 \cdot \text{iq} \\
\text{educ} \cdot (\ln(\text{earn}) - \beta_0 - \beta_1 \cdot \text{educ} - \beta_2 \cdot \text{exper} - \beta_3 \cdot \text{exper}^2 - \beta_4 \cdot \text{iq}) \\
\text{eper} \cdot (\ln \text{earn}) - \beta_0 - \beta_1 \cdot \text{educ} - \beta_2 \cdot \text{exper} - \beta_3 \cdot \text{exper}^2 - \beta_4 \cdot \text{iq}) \\
\text{exper}^2 \cdot (\ln(\text{earn}) - \beta_0 - \beta_1 \cdot \text{educ} - \beta_2 \cdot \text{exper} - \beta_3 \cdot \text{exper}^2 - \beta_4 \cdot \text{iq}) \\
\text{iq} \cdot (\ln(\text{earn}) - \beta_0 - \beta_1 \cdot \text{educ} - \beta_2 \cdot \text{exper} - \beta_3 \cdot \text{exper}^2 - \beta_4 \cdot \text{iq}) \\
\ln(\text{earn}) - 2.0647 \\
\ln(\text{earn})^2 - 4.5258 \\
\text{educ} - 13.97 \\
\ln(\text{earn}) \cdot \text{educ} - 29.099
\end{pmatrix}.
$$

(a) Estimate β by least squares using only the NLS data (and thus only the first five moments).

(b) Use the ols estimates as initial consistent estimates to consistently estimate the optimal weight matrix. Note: you can estimate the optimal weight matrix in different ways given initial estimates for the parameters, for example by demeaning the moments first. To ensure that everybody is estimating exactly the same thing, estimate the weight matrix given initial estimates $\tilde{\theta}$ as

$$
\hat{\Delta}^{-1} = \left[\frac{1}{N} \sum_{i=1}^{N} \psi(z_i, \tilde{\theta}) \cdot \psi(z_i, \tilde{\theta})' \right]^{-1} .
$$

(c) Use DFP to find the optimum. You can incorporate parts of the handout on the webpage in your program. Report the values of the parameters for each iteration. (If you need more than 50 iterations, just report the values every tenth iteration. Something must have gone wrong!) How do the efficient estimates differ from the ols ones?

(d) Calculate the standard errors for the gmm estimates. How do they compare to the ols standard errors?

(e) Use the objective function to test whether the overidentifying restrictions are satisfied. What are the degrees of freedom?