**Econometrics**

**PhD in Business Administration and Quantitative Methods**

**Course 2006-2007 - June 21th, 2007**

Name:

Please, answer 5 of the following 6 questions.

(1) Consider the following regression model:

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

where the $x_i$ are non-stochastic and the $\{\varepsilon_i\}$ are *i.i.d.* with $E[\varepsilon|x] = 0$ and $E[\varepsilon^2|x] = \sigma^2$.

a) What is the minimum mean squared error linear estimator of $\beta$?

b) For the estimator in part (a), $\widetilde{\beta}$, show that the ratio of the mean squared error of $\widetilde{\beta}$ to that of the ordinary least squared estimator, $\widehat{\beta}$, is

$$\frac{MSE(\widetilde{\beta})}{MSE(\widehat{\beta})} = \frac{\tau^2}{1+\tau^2}, \text{ where } \tau^2 = \frac{\beta^2}{\sigma^2/\boldsymbol{x}'\boldsymbol{x}}.$$

c) Note that $\tau$ is the square of the population analog to the "t ratio" for testing the hypothesis that $\beta = 0$. How do you interpret the the behavior of this ratio $\tau$?

(2) Consider the following regression model:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \alpha\varepsilon_i,$$

where the $x_i$ are non-stochastic.

a) Assume that the $\{\varepsilon_i\}$ are *i.i.d.* Cauchy. So,

$$f_{\varepsilon_i}(x) = \frac{1}{\pi(1+x^2)}, \ -\infty < x < +\infty.$$

The Cauchy density has a shape similar to a normal density, but with much thicker tails. Thus, extremely small and large errors occur much more frequently with this density than would happen if the errors were normally distributed. Find the score function $g_n(\theta)$ where $\theta = (\boldsymbol{\beta}', \alpha)'$.

b) Assume that the $\{\varepsilon_i\}$ are *i.i.d.* $\mathcal{N}(0, \sigma^2)$. Find the score function $g_n(\theta)$ where $\theta = (\boldsymbol{\beta}', \alpha)'$.

c) Compare the first order conditions that define the maximum likelihood estimators of parts (a) and (b) and interpret the differences. Why are the first order conditions that define an efficient estimator different in the two cases?

(3) Consider the following regression model:

$$y_i = \mu + \varepsilon_i,$$

where $E[\varepsilon_i|x_i] = 0$, $Cov(\varepsilon_i, \varepsilon_j|x_i, x_j) = 0$ for $i \neq j$, but $E[\varepsilon_i^2|x_i] = \sigma^2 x_i^2$ and $x_i > 0$.

a) Given a sample of observation on $y_i$ and $x_i$, what is the most efficient estimator of $\mu$? What is its variance?

b) What is the OLS estimator of $\mu$, and what is the variance of the ordinary least squares estimator?

c) Prove that the estimator in part (a) is at least as efficient as the estimator in part (b).

d) Suppose that the $\{\varepsilon_i\}$ are normally distributed, with mean zero and variance $\sigma^2(1+(\gamma x)^2)$. Show that $\sigma^2$ and $\gamma^2$ can be consistently estimated by a regression of the least squares residuals on a constant and $x^2$.

(4) Consider the following regression model:

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon},$$

where $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ have $K_1$ and $K_2$ columns respectively. The restriction is $\boldsymbol{\beta}_2 = \boldsymbol{0}$.

a) Prove that the restricted estimator is simply $(\widehat{\boldsymbol{\beta}}_1^*, \boldsymbol{0})$, where $\widehat{\boldsymbol{\beta}}_1^*$ is the least squares estimator in the regression of $\boldsymbol{y}$ on $\boldsymbol{X}_1$.

b) Prove that if the restriction is $\boldsymbol{\beta}_2 = \boldsymbol{\beta}_2^0$ for a nonzero $\boldsymbol{\beta}_2^0$ then the restricted estimator for $\boldsymbol{\beta}_1$ is

$$\widehat{\boldsymbol{\beta}}_1^* = (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'(\boldsymbol{y} - \boldsymbol{X}_2\boldsymbol{\beta}_2^0).$$

c) Prove the result that the $R^2$ associated with a restricted least squares estimator is never larger than that associated with the unrestricted least squares estimator.

(5) Consider the following system of simultaneous equations:

$$\begin{aligned} y_{t1} &= \gamma_{31}y_{t3} + \beta_{11}x_{t1} + \beta_{21}x_{t2} + \varepsilon_{t1}, \\ y_{t2} &= \gamma_{32}y_{t3} + \beta_{32}x_{t3} + \varepsilon_{t2}, \\ y_{t3} &= \gamma_{23}y_{t2} + \varepsilon_{t3}, \end{aligned}$$

where $y$ and $x$ indicate dependent and predetermined variables, respectively.

a) Estimate the second equation with the 2SLS method using the data:

$$\begin{array}{lll} \sum \widehat{y}_{t3}y_{t3} = 100, & \sum \widehat{y}_{t3}y_{t2} = 5, & \sum \widehat{y}_{t2}y_{t3} = 6, \\ \sum x_{t3}^2 = 2, & \sum x_{t3}\widehat{y}_{t2} = 11, & \sum x_{t2}\widehat{y}_{t2} = 12, \\ \sum \widehat{y}_{t3}x_{t3} = -10, & \sum x_{t3}y_{t2} = 15 & \sum x_{t2}y_{t2}, = 10. \end{array}$$

Here, $\widehat{y}$ denotes the estimate of $y$ resulting from the estimated reduced form of the model.

b) Derive the reduced form for the second equation.

c) Give conditions under which the OLS estimator of the second structural equation is consistent.

(6) Consider the following regression model:

$$y_t = \beta y_{t-1} + \varepsilon_t, \ \varepsilon_t = \upsilon_t + \rho\upsilon_{t-1}, \ t = 1, 2, \ldots, T,$$

with $|\beta| < 1$, $|\rho| < 1$, $\upsilon_t \sim i.i.d.(0, \sigma^2)$ and $E[\upsilon_t y_{t-s}] = 0$ for $s \geq 1$.

a) Show that the OLS estimator of $\beta$ fulfills

$$P\lim \widehat{\beta} = \beta + \rho\sigma^2/\tau^2, \text{ where } \tau^2 = Var(y_t).$$

b) Propose a suitable instrumental variable for $y_{t-1}$ and argue that your instrumental variable fulfills the conditions for a consistent IV–estimation under the present model assumption.