New York University | **Spring 2003**
Econometrics II | **Professor: Giorgio Topa**
G31.2101 | **TA: Luca Flabbi**

PROBLEM SET 4

due Friday May 2, 2003

1. Consider estimation of the following panel data model

$$\begin{aligned} y_{it} &= \alpha + x_{it}\beta + u_{it}, \quad i = 1, ..., I;\ t = 1, 2; \\ u_{it} &= \eta_i + \varepsilon_{it}, \end{aligned}$$

where

$$\begin{aligned} E(\varepsilon_{it}|x_i, \eta_i) &= 0, \quad \forall\, i, t \\ E(\varepsilon_{it}^2|x_i, \eta_i) &= \tau x_{it}^2 \quad \forall\, i, t \\ E(\varepsilon_{i1}\varepsilon_{i2}|x_i, \eta_i) &= 0, \quad \forall\, i, t \\ E(\eta_i|x_i, \varepsilon_{it}) &= 0, \quad \forall\, i, t \\ E(\eta_i^2|x_i, \varepsilon_{it}) &= \sigma_\eta^2 \quad \forall\, i, t. \end{aligned}$$

Here x_{it} is a scalar, and x_i denotes $(x_{i1}\, x_{i2})'$. It is assumed that the variable x_{it} always takes positive values.

(a) Define an efficient FGLS estimator of α and β if one exists.

(b) Suppose that instead of assuming $E(\eta_i|x_i, \varepsilon_{it}) = 0$ and $E(\eta_i^2|x_i, \varepsilon_{it}) = \sigma_\eta^2\ \forall\, i, t$, we only assume that

$$E(\eta_i|x_i, \varepsilon_{it}) = E(\eta_i|x_i) \neq 0 \quad \forall\, i, t.$$

Is the 'first-difference' OLS estimator consistent in this case? If so, can you define a more efficient FGLS estimator that is also consistent? If so, please describe it.

2. **GMM estimators for Linear Models**

In this question you are asked to work out results for GMM estimators for a single equation linear model. The idea is to label alternative estimators by the choice of moment conditions used in estimation. Consider the following model:

$$y_t = x_t'\beta_0 + \varepsilon_t, \qquad E(z_t\varepsilon_t) = 0$$

where y_t is a scalar, x_t is a $(k \times 1)$ vector, z_t is a $(m \times 1)$ vector of instrumental variables, with $m > k$. Assume that (y_t, x_t, z_t) is stationary, that $E(z_t x_t')$ has full rank k and that $\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t\varepsilon_t\right)$ converges in distribution to a Normally distributed random vector with mean zero and covariance matrix V_0.

(a) Let A_0 denote a $(k \times m)$ matrix of real numbers that selects the moment conditions to be used in estimation. In other words, the estimator b_T associated with A_0 satisfies

$$A_0 \frac{1}{T}\sum_{t=1}^{T} z_t(y_t - x_t' b_T) = 0$$

Show that if $A_0 E(z_t x_t')$ has full rank k, the estimator b_T converges almost surely to β_0. Show that $\sqrt{T}(b_T - \beta_0)$ converges in distribution to a normally distributed random vector with mean zero. Provide an expression for the asymptotic covariance matrix.

(b) Let $\widetilde{A}$ be given by

$$\widetilde{A} = CE(x_t z_t')(V_0)^{-1}$$

for some non-singular matrix C. Derive an expression for the asymptotic covariance matrix of an estimator that uses selection matrix $\widetilde{A}$. Does this covariance matrix depend on the matrix C, why or why not? Show that $\widetilde{A}$ is the most "efficient" choice of A_0.

(c) Suppose that the selection matrix A_0 is replaced by A_T where $A_T \xrightarrow{a.s.} A_0$. Verify that the results in part (a) are still true.

3. **Conditional Homoskedasticity and Two-Stage Least Squares**

In addition to the assumptions made in question 1, suppose that

$$E(\varepsilon_t | z_t, z_{t-1}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = 0.$$

(a) Provide a characterization of V_0.

(b) Add the conditional homoskedasticity assumption:

$$E(\varepsilon_t^2 | z_t, z_{t-1}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}, ...) = \sigma^2,$$

and show how to use this assumption to provide an alternative characterization of V_0 that involves only the second moments of z_t and ε_t. Show that one version of $\widetilde{A}$ in part (b) of Question 1 is

$$\widetilde{A} = E(x_t z_t')(E z_t z_t')^{-1}.$$

Notice that this "efficient" selection matrix does not depend directly on the error terms ε_t. Suggest a way of consistently estimating this efficient selection matrix and provide an expression for the resulting asymptotic covariance matrix of the estimator b_T.

(c) The estimator you just developed is the Two Stage Least Squares estimator. Is the 2SLS estimator asymptotically efficient when the conditional homoskedasticity assumption in part (b) is relaxed? Why or why not?