BIRKBECK
(University of London)

MSc EXAMINATION FOR INTERNAL STUDENTS

Department of Economics, Mathematics, and Statistics

ECONOMETRICS/ ECONOMETRICS for PG Certificate (Part 2)

EMEC026S7/BUEM007H7

Friday, 2 June 2017, 10.00 am - 12.10 pm (includes 10 minutes reading time).

Answer **ANY TWO** questions. All questions carry the same weight; the relative weight of sub-questions is indicated in square brackets. Any results used in lectures can be used without proof, but need to be properly stated.

---

1. Consider the following system of equations,

$$\begin{aligned} y_n &= x_n\beta + \epsilon_n \\ x_n &= z_n\gamma + u_n, \quad n = 1, \cdots, N, \end{aligned}$$

where both x_n and z_n are scalar regressors, and $\gamma \neq 0$. Suppose that

$$\begin{aligned} \mathbb{E}[x_n\epsilon_n] &\neq 0 \\ \mathbb{E}[z_n\epsilon_n] &= 0 \\ \mathbb{E}[z_n u_n] &= 0. \end{aligned}$$

(a) [25%] Derive the reduced form equation [5%]

$$y_n = \lambda z_n + \nu_n,$$

and show that [each 5%]

$$\begin{aligned} \beta &= \lambda/\gamma \\ \mathbb{E}[z_n\nu_n] &= 0 \\ \mathrm{var}(\nu_n) &= \beta^2\sigma_u^2 + \sigma_\epsilon^2, \end{aligned}$$

where σ_u^2 and σ_ϵ^2 are the variances of u_n and ϵ_n, respectively.

Also, derive the instrumental variable estimator for β in this problem [5%].

(b) [25%] Suppose γ is estimated by OLS, regressing the vector $\mathbf{X}' = [x_n, n = 1, \cdots, N]$ on the vector $\mathbf{Z}'$; and that λ is estimated by OLS, regressing $\mathbf{Y}$ on $\mathbf{Z}$. Show how to use these two OLS estimates to obtain a consistent estimator for β [10%]. Be sure to provide an argument supporting the consistency of the two OLS estimators [each 5%] and your proposed estimator of β [5%].

(c) [25%] Let $\theta' = (\lambda, \gamma)$ and $\xi_n' = (\nu_n, u_n)$. Derive an expression for $\sqrt{N}\left(\hat{\theta} - \theta\right)$ that is a function of ξ_n.

(d) [25%] Suppose that $\frac{1}{N}(\sum_{n=1}^N z_n^2) \xrightarrow{p} \sigma_z^2$ as $N \to \infty$, where $\sigma_z^2 = \mathbb{E}[z_n^2]$ for all n. And define $\Omega_\xi = \mathbb{E}[z_n^2\xi_n\xi_n']$. Derive the joint asymptotic distribution of $\sqrt{N}\left(\hat{\theta} - \theta\right)$ as $N \to \infty$ [15%]. Are $\hat{\lambda}$ and $\hat{\gamma}$ asymptotically independent? Why, or why not? [10%]

2. Suppose a researcher is interested in estimating the effect of a health care initiative (e.g. free pneumonia vaccination) aimed at the elderly in California (say, people above the age of 70) on a health outcome, measured by the variable Y. The researcher collected data on Y for a random sample from the population of elderly Californian residents before and after the introduction of the health care initiative, and also for a random sample of the same size from the population of elderly residents in neighbouring Oregon where no such health initiative was implemented.

   (a) [25%] Describe how the data can be used to identify the effect of the Californian policy initiative on the health outcome. Be sure to state any additional assumptions you need.

   (b) [25%] Residents of California have a higher per capita income than residents of Oregon. Since health outcomes are often found to be correlated with income, the researcher is concerned that observations in the treatment sample (elderly in California) are subject to a different inter-temporal change in health outcomes than those in the control sample. If this were indeed the case, would this bias the treatment effect estimate of the approach in (a), and if so how?

   (c) [25%] To account for the potential problem in (b), the researcher also collects data on other Californian and Oregon residents, whose age is below 70, for the two time periods, before and after the inctroduction of the health care initiative. Describe how your model in (a) can be extended to also capture these additional data.

   (d) [25%] Explain how the data, within the model constructed in (c), identifies the effect of the California health care initiative in the presence of the concerns raised in (b). Be sure to state any additional assumptions you may need.

3. Consider the regression model

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \quad 1 = 1, \cdots, N; t = 1, \cdots, T,$$

where x_{it} is a scalar regressor and α and β are scalar parameters.

(a) [25%] Let $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\bar{\bar{y}} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$, and analogously for $\bar{x}_i, \bar{u}_i, \bar{\bar{x}}$ and $\bar{\bar{u}}$. Show that

$$y_{it} - \bar{\bar{y}} = \beta\left(x_{it} - \bar{x}_i\right) + \beta\left(\bar{x}_i - \bar{\bar{x}}\right) + u_{it} - \bar{\bar{u}}.$$

(b) [25%] Suppose OLS is used to regress the $NT \times 1$ vector $(y_{it} - \bar{\bar{y}}, i = 1, \cdots, N, t = 1, \cdots, T)$ on the $NT \times 2$ matrix $\mathbf{X}$ whose first column is the vector $\mathbf{X}'_{.1} = (x_{it} - \bar{x}_i, i = 1 \cdots, N, t = 1 \cdots, T)$ and whose second column is the vector $\mathbf{X}'_{.2} = (\iota_T(\bar{x}_i - \bar{\bar{x}})', i = 1, \cdots, N)$, where ι_T is a $T \times 1$ vector of 1s; i.e. the regression is

$$y_{it} - \bar{\bar{y}} = \beta_1\left(x_{it} - \bar{x}_i\right) + \beta_2\left(\bar{x}_i - \bar{\bar{x}}\right) + u_{it} - \bar{\bar{u}}.$$

Derive a closed form expression for the OLS estimator of the coefficient on β_1, and a closed form solution for the OLS estimator of the coefficient on β_2 in this regression.

(c) [25%] Suppose $u_{it} = \mu_i + \epsilon_{it}$, where the random variables μ_i are i.i.d. with mean zero and variance σ_μ^2, and the random variables ϵ_{it} are i.i.d. with mean zero and variance σ_ϵ^2, and μ_i and ϵ_{it} are independent across all i and t. Derive the respective variance of the within and between estimators of β in the *original* model under these assumptions and explain how you would assess the efficiency of the within relative to the between estimator.

(d) [25%] Show that under the assumptions in (c) the OLS and GLS estimators for the regression in (b) are equivalent for large NT.

End of Examination Paper