Advanced Micro Econometrics TI 152 Fall 2020

Computer assignment

The due date for this assignment is Tuesday November 24-th. The total length of your assignment should be at most four pages (including figures). There is no need to hand in any computer code.

1. Size distortions: To illustrate the size distortions of the 2SLS t-statistic, we simulate data from the following model:

$$
\begin{aligned}
Y &= X\beta + \varepsilon \\
X &= Z\Pi + V
\end{aligned}
$$

   where Y and X are $n \times 1$ vectors which contain the endogenous variables and Z is a $n \times k$ matrix of instruments. ε and v are $n \times 1$ vectors that contain the disturbances. The different rows of $(\varepsilon \vdots V)$, $\binom{\varepsilon_i}{V_i}'$, $i = 1, \ldots, n$, are independently normal distributed: $\binom{\varepsilon_i}{V_i} \sim N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 1 & \vdots & \rho \\ \rho & & 1 \end{pmatrix}$. We use $n = 100$, $k = 10$, $\Pi = a \times e_{10}$ with e_{10} a 10×1 vector whose top element is one and all remaining elements are equal to zero. All elements of Z are independently standard normal distributed. We only simulate them once and keep them fixed throughout the simulation experiment.

   We use seven values of a : (0.3 0.25 0.2 0.15 0.1 0.05 0) and ten different values of ρ : (0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.9 0.95).

   For each value of a make a figure of the rejection frequency as a function of ρ when testing $\mathrm{H}_0 : \beta = 0$ with 95% significance using the 2SLS t-statistic (so five figures which show the rejection frequency as a function of a). Use 5000 simulations from the above model.

   What do you conclude?

2. Compute and make a figure of the 95% critical value function of the LR statistic as a function of $r(\beta_0)$ for $k = 10$. What can you say about the critical value when $r(\beta_0) = 0$ or infinite?

3. Repeat the excercise in 1 for the AR, score and LR statistics. What do you conclude?

4. Compute and make a figure of the 95% critical value function of the LR statistic as a function of $r(\beta_0)$ for $k = 4$.

   Question 5 is on the other side.

5. Card (1993)[1] analyzes the return on education. He uses different proximity to college variables as instruments. The file assigmentweakinstruments.mat contains that part of the Card data which we use for this assignment. The different variables in the file are: nearc2: if near a 2 year college, nearc4: if near a 4 year college, nearc4a: if near a 4 year community college, nearc4b: if near a 4 year private college, ed: years of education, wage: log-earnings, age: age in years, age2: squared age, exper: experience, exper2: experience squared, south: lives in the South, smsa: lives in a metropolitan area, race: racial indicator.

   The variables wage and ed constitute the endogenous variables (y and x), nearc2, nearc4, nearc4a, nearc4b are instruments (z) and exper, exper2, south, smsa, race and the constant term are the included exogenous variables (w) (We do not use age and age2).

   (a) Using only nearc2 as an instrument, construct the 95% confidence set for the return on education using the 2SLS t-statistic and the AR statistic.

   (b) Is there a difference between these confidence sets and if so can you explain why this difference occurs?

   (c) What is the value of the first stage F-statistic and what does the value of the AR statistic look like when the tested parameter is large.

   (d) We did not use the LM and LR statistics in a or did we?

   (e) Using nearc4, nearc2, nearc4a and nearc4b as instruments, construct the 95% confidence set for the return on education using the 2SLS t-statistic, AR, LM and LR statistics.

   (f) Is there a difference between these confidence sets and if so can you explain why this difference occurs?

[1] Card, D. Using geographic variation in college proximity to estimate the return to schooling. In L.N. Christofides, E.K. Grant and R. Swidinsky, editor, Aspects of Labour Market Behaviour: essays in honor of John Vanderkamp, pages 201—222. University of Toronto Press, Toronto, Canada, 1995. (NBER Working Paper 4483 (1993)).