UCLA, Department of Economics, Patrik Guggenberger
Economics 203C (Spring, 2009), April 30, 2009

Problem Set V (due Wed 5/13/09)

1. Consider the linear simultaneous equations model

$$\begin{aligned} Y &= X\beta + \varepsilon, \\ X &= Z\Pi + v, \end{aligned} \tag{1}$$

where Y is $n \times 1$, X is $n \times d$. The observations are i.i.d. and you can assume conditional homoskedasticity. The sample size n is larger than d. Interest focuses on estimation of β.

i) If Z is an $n \times n$ invertible matrix, show that 2SLS and OLS are numerically identical.

Now consider 2SLS estimators $\hat{\beta}_i$ $(i = 1, 2)$ based on instruments Z_i of dimensions $n \times k_i$, where $k_1 < k_2$ and Z_1 consists of the first k_1 columns of Z_2.

ii) Verify that the asymptotic variance of $\hat{\beta}_2$ is smaller than the one of $\hat{\beta}_1$ (in the positive definite sense).

iii) Discuss the statement "The more instruments we use for 2SLS, the better the estimator becomes".

2. Investigate the finite sample properties of OLS and the least absolute deviation estimator (LAD) in the following Monte Carlo setup. Consider the model

$$y_i = x_i'\beta + u_i,\ i = 1, ..., n \tag{2}$$

where $\beta \in R^k$ and the first component of x_i is the constant 1. The errors u_i are identically and independently (iid) distributed.

Consider three different setups. Setup I has $u_i \sim iid\ N(0, 1)$, in Setup II u_i has a t-distribution with 3 degrees of freedom, and in Setup III u_i is iid Cauchy.

In both setups, the components 2 to k of x_i are taken as $iid\ N(0, 1)$. Compare the finite sample bias, variance, and mean squared error of $\widehat{\beta}_2$ where $\widehat{\beta} = (\widehat{\beta}_1, ..., \widehat{\beta}_k)'$ and $\widehat{\beta}$ is either the OLS estimator or the LAD estimator, where the latter is defined as

$$\arg\min_{\beta} \textstyle\sum_{i=1}^{n} |y_i - x_i'\beta| \tag{3}$$

How do β, k, n affect the results? (Take as one scenario $\beta = (0, ..., 0)' \in R^k$, $k = 2$, and $n = 100$ and do "comparative statics" by looking at other parameter choices) How do the different distributions of u_i affect the results?

3. True, questionable or false? Just stating t,q or f does not earn you any points. The explanation is what counts.

(i) One should not do inference based on a Wald test when instruments are potentially weak because then the power of the test can be very low.

(ii) In a linear regression model with omitted regressors, OLS estimation is always inconsistent.

(iii) In the linear IV model estimated by GMM let A_n be an optimal weight matrix under the assumption of conditional homoskedasticity. If we use GMM with weight matrix A_n in a scenario where we have conditional heteroskedasticity, the estimator is typically inconsistent.

(iv) A possible advantage of the likelihood ratio statistic over the Wald statistic is its compact formula and easy implementation.

(v) If $1/X_n = O_p(1)$ then it can not be the case that $X_n = o_p(1)$.

(vi) If available, one should always include an IQ score as a regressor in a wage regression to overcome the inconsistency problem of OLS caused by unobserved ability.

4. a) Let $\{X_t\}$ for $t \in \mathbb{Z}$ be the stationary solution of the non-causal AR(1) equations,

$$X_t = \phi X_{t-1} + Z_t,\ \{Z_t\} \sim WN(0, \sigma^2),\ |\phi| > 1. \tag{4}$$

Show that $\{X_t\}$ also satisfies the causal AR(1) equations,

$$X_t = \phi^{-1} X_{t-1} + \widetilde{Z}_t,\ \{\widetilde{Z}_t\} \sim WN(0, \widetilde{\sigma}^2), \tag{5}$$

for a suitably chosen white noise process $\{\widetilde{Z}_t\}$. Determine $\widetilde{\sigma}^2$. (Recall the formula $\sum\limits_{i=0}^{k} q^i = \frac{1-q^{k+1}}{1-q}$ for $|q| < 1$.)

b) Let $\{X_t\}$ and $\{Y_t\}$ for $t \in \mathbb{Z}$ be independent stationary $AR(p)$ processes. Show that $Z_t := X_t + Y_t$ is $ARMA(2p, p)$. Derive the corresponding result for $MA(q)$ processes. (See Hamilton (1994, chapter 4.7)).