Advanced Econometrics II
Homework 3

Deadline: 2011-01-31 09:30

You are to hand in your homework via email to ta@zamojski.net by 9:30 on Tuesday. I would appreciate if your answers were TeXed in full, if you insist you can scan your handwritten answers at TI and send them to me as well.

You can work in groups of two.

If there is a computational exercise included in the homework, the quality of your coding will be judged (commenting, efficiency, etc.). Your code should be easy to read. Things that make it easier: lots of comments (e.g. explaining what loops are meant to do), camel case variable names (e.g. mErrorsUniformlyDistributed), consistency. You are to include your code in the body of the report, e.g. if you are TeXing your answers then with the listings package. You can use Ox, Python, C++, or Matlab. If I am not able to run your code after extracting your answers (assuming they are zipped) to a separate folder you will lose points. Looking ahead, it is in your best interest to combine Ox (computations, sometimes graphics) and Python (database management, multiprocessing management, graphics) as it will cut the time needed for simulations considerably. In the empirical exercises you are expected to provide comments for your results and methods (e.g tests) used.

Exercise 1

Consider a GMM criterion function of the form:

$$argmin_{\beta} \left\{ g^T \left(\beta \right) V^{-1} g \left(\beta \right) \right\} \tag{1}$$

Assume that V^{-1} can be decomposed into ZZ^T where Z is a lower triangular matrix. Then the new criterion yields:

$$argmin_{\beta} \left\{ g^T \left(\beta \right) ZZ^T g \left(\beta \right) \right\} \tag{2}$$

1. Map this into notation of the book. How general is this case?
2. What is the importance and interpretation of Z^T?
3. Interpret the GMM procedure when:

$$Z^T = \begin{bmatrix} x & y \\ 0 & z \end{bmatrix} \tag{3}$$

4. Can you conceive of a situation where you would prefer to restrict elements of V^{-1} (i.e. of Z^T)? You may illustrate with x, y, z from the previous sub-question. Will this affect consistency of your estimator and/or cause problems for testing?

Exercise 2

For some r, z, and K asymptotic variance of the GMM estimator is

$$(rz)^{-1} rKr^T \left((rz)^T\right)^{-1} \tag{4}$$

By proposing and deriving the difference between the two, prove that for some choice of $r = \tilde{r}$ this variance can be minimized. Interpret your suggestion of $\tilde{r}$. For ease of exposition you may cast this exercise in the notation of the book. You may use the result of Exercise 3.8 without deriving it.

Exercise 3

1. In what sense is the *efficient* GMM efficient and how it relates to moment conditions?
2. Consider over-identified estimation based on the moment conditions (9.31)

$$\mathrm{E}\left[\mathbf{W}^{\mathrm{T}} (\boldsymbol{\Omega})^{-1} (\mathbf{y} - \mathbf{X}\beta)\right] = 0 \tag{5}$$

   You may assume that the instruments $\mathbf{W}$ satisfies the predeterminedness condition. Derive the GMM criterion function for these theoretical moment conditions, and report the estimating equations that result from this procedure.

3. Suppose that $S\left(\bar{\mathbf{X}}\right)$, the span of the $n \times k$ matrix $\bar{\mathbf{X}}$ of optimal instruments, is a linear subspace of $S(\mathbf{W})$, the span of the transformed instruments. Show that, in this case, the estimating equations are asymptotically equivalent to

$$\bar{\mathbf{X}}^{\mathrm{T}} (\boldsymbol{\Omega})^{-1} (\mathbf{y} - \mathbf{X}\beta) = 0 \tag{6}$$

4. Derive the asymptotic covariance matrix of the estimator in 2., and show that it is 'bigger' than the asymptotic covariance matrix of the efficient estimator.

Exercise 4

Assume serial independence and homoskedasticity of error terms and consider the criterion function:

$$(\mathbf{y} - \mathbf{X}\beta)^{\mathrm{T}} \mathbf{P}_{\mathbf{W}} (\mathbf{y} - \mathbf{X}\beta) \tag{7}$$

1. How is the efficiency and consistency of the resulting GMM estimator affected by substitution of $\left(\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)^{-1}$ by some asymptotically non-random, symmetric, and positive definite matrix $\mathbf{A}$?
2. To show difference in efficiency – instead of comparing precision matrices – it is enough to show that $\hat{\beta}_{(\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}}$ is asymptotically uncorrelated with $\hat{\beta}_{\mathbf{A}} - \hat{\beta}_{(\mathbf{W}^{\mathrm{T}}\mathbf{W})^{-1}}$. Explain why and derive this result.