Econ 220B Final Exam
Winter 2014

DIRECTIONS: No books or notes of any kind are allowed. 250 points are possible on this exam.

1.) (30 points total) An eager undergraduate student has just learned about the fiscal multiplier, which claims that if real government spending G_t were to increase by one dollar, real GDP y_t would increase by more than one dollar. The student proposes to test this using a time series regression

$$y_t = \alpha + \beta G_t + u_t$$

using annual data for the United States, $t = 1947$ to 2012.

a.) (15 points) Is serial correlation a potential concern about the above regression? If so, suggest a simple test the student might perform to see whether this is an issue and a possible way to fix the problem.

b.) (15 points) Is correlation between u_t and G_t a potential concern about the above regression? If so, identify what you feel is the single most important reason, and explain to the student whether this is likely to bias the OLS estimate β up or down.

2.) (40 points total) Consider a regression model in which the errors u_t are independent across observations and heteroskedastic:

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + u_t$$

$$E(u_t|\mathbf{X}) = 0$$

$$E(u_t u_s|\mathbf{X}) = 0 \quad \text{for } t \neq s$$

$$E(u_t^2|\mathbf{X}) = \alpha_0 + \alpha_1 x_{t1}.$$

Here $\mathbf{X}$ denotes the full $(T \times k)$ matrix of observations on the $(k \times 1)$ vector $\mathbf{x}_t$, x_{t1} denotes the first element of $\mathbf{x}_t$, and α_0 and α_1 are unknown parameters.

a.) (15 points) Describe a quick test of the null hypothesis $\alpha_1 = 0$ that you could perform by looking at $T \cdot R^2$ where R^2 denotes the coefficient of determination of a simple regression.

b.) (10 points) Give a formula for the variance of the OLS estimate $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ that would be asymptotically valid even if $\alpha_1 \neq 0$.

c.) (15 points) Describe how you would estimate $\boldsymbol{\beta}$ using feasible GLS.

3.) (100 points total) If $E[\mathbf{h}(\boldsymbol{\theta}_0, \mathbf{w}_t)] = \mathbf{0}$, the GMM estimate of $\boldsymbol{\theta}$ is defined as the value that minimizes

$$T[\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)]'\hat{\mathbf{S}}^{-1}[\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)]$$

where

$$\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T) = T^{-1}\sum_{t=1}^{T}\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)$$

and $\hat{\mathbf{S}}$ is a consistent estimate of

$$\mathbf{S} = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\sum_{v=-\infty}^{\infty} E\left\{[\mathbf{h}(\boldsymbol{\theta}_0, \mathbf{w}_t)]\left[\mathbf{h}(\boldsymbol{\theta}_0, \mathbf{w}_{t-v})\right]'\right\}.$$

The usual asymptotic distribution of the GMM estimator is

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{V})$$

where $\mathbf{V}$ can be estimated consistently from

$$\hat{\mathbf{V}} = (\hat{\mathbf{D}}\hat{\mathbf{S}}^{-1}\hat{\mathbf{D}}')^{-1}$$

$$\hat{\mathbf{D}}' = \left.\frac{\partial \mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)}{\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.$$

In this question you are asked to use the GMM framework to show how instrumental variables could be used to estimate a nonlinear regression model.

Suppose

$$y_t = q(z_t, \theta) + u_t$$

where θ is a scalar parameter you are trying to estimate from observations on the scalars y_t and z_t along with an observed $(m \times 1)$ vector of instruments $\mathbf{x}_t$ satisfying $E\{\mathbf{x}_t[y_t - q(z_t, \theta_0)]\} = \mathbf{0}$ for θ_0 the true value of θ. Suppose also that $\mathbf{x}_t u_t$ is a stationary and ergodic martingale difference sequence with a known variance-covariance matrix given by $E(u_t^2 \mathbf{x}_t \mathbf{x}_t') = \sigma^2 \mathbf{Q}$ where σ^2 is a known scalar and $\mathbf{Q}$ is a known positive definite $(m \times m)$ matrix. Note that the econometrician already knows the values of σ^2 and $\mathbf{Q}$ and therefore does not need to try to estimate σ^2 or $\mathbf{Q}$ from the data. We have a sample of size T on the variables, that is, we observe $\{y_1, z_1, \mathbf{x}_1, ..., y_T, z_T, \mathbf{x}_T\}$.

a.) (20 points) Calculate the expressions for the functions $\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)$ and $\mathbf{g}(\boldsymbol{\theta}; \mathbf{Y}_T)$ for this particular example.

b.) (20 points) Find the first-order condition for the GMM estimate of $\hat{\theta}$. Hint: check the dimensions of any vectors or matrices that appear in your answer to make sure your answer makes sense.

c.) (10 points) Find a closed-form expression for $\hat{\theta}$ in the special case when $q(z_t, \theta) = z_t \theta$.

d.) (20 points) What does it mean for the instruments $\mathbf{x}_t$ to be relevant for this example? Hint: if you're not sure how to answer this, see if you can at least give the answer for the special case when $q(z_t, \theta) = z_t \theta$.

e.) (15 points) Give an expression for the GMM t test of the null hypothesis $\theta = 0$.

f.) (15 points) Give an expression for an alternative test you could use of the null hypothesis $\theta = 0$ that would be have the correct size even if instruments are weak.

4.) (80 points total) In this problem you are are asked to analyze maximum likelihood estimation and quasi-maximum likelihood estimation for an i.i.d. sample of Normal variables with known variance and unknown mean. Hint: if you have trouble remembering particular formulas, refer back to the general results about GMM estimation that were provided in the statement at the beginning of question 3 above.

Suppose $y_t \sim$ i.i.d. $N(\mu, 1)$ so that the log likelihood function for $\mathbf{Y}_T = (y_1, y_2, ..., y_T)'$ is

$$\ell(\mu; \mathbf{Y}_T) = -(T/2)\log 2\pi - (1/2)\sum_{t=1}^{T}(y_t - \mu)^2.$$

a.) (20 points) Find the score associated with observation t and show directly that this follows a martingale difference sequence if the model is correctly specified.

b.) (10 points) Find $\hat{\mu}$, the maximum likelihood estimate of μ for this example.

c.) (10 points) In this case there is only a single unknown parameter (μ) so the information matrix is a scalar. Calculate an estimate of the information matrix based on the second derivative of the log likelihood function and use this to calculate an asymptotic standard error for the MLE $\hat{\mu}$.

d.) (15 points) Calculate an alternative estimate of the information matrix based on the average outer product of the scores and show that its plim is the same as the plim of the estimator in (c).

e.) (10 points) If a researcher were to give a quasi-maximum likelihood interpretation to $\hat{\mu}$, what would it be? Give a concise mathematical definition of the "pseudo-true" value μ_0 which one would claim to estimate by quasi-maximum likelihood.

f.) (15 points) Suppose you believed that the data y_t were serially uncorrelated but not necessarily Gaussian or i.i.d. Find the expression for the quasi-MLE standard error for $\hat{\mu}$. In what sense is this any more robust than the standard error you calculated in (c)?