Econ 220B Final Exam
Winter 2013

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

1.) (90 points total) A researcher is interested in the relation between someone's income and how much their father earns. Consider the following definitions of the observed variables

$y_t =$ income of person t

$\overline{y} = T^{-1}\sum_{t=1}^{T} y_t =$ average income of all people in sample

$\tilde{y}_t = y_t - \overline{y} =$ deviation of income of person t from the sample mean

$x_t =$ income of father of person t

$\overline{x} = T^{-1}\sum_{t=1}^{T} x_t =$ average income of all the fathers in sample

$\tilde{x}_t = x_t - \overline{x} =$ deviation of person t's father income from the sample mean.

$T = 66 =$ sample size

An ordinary least squares regression of $\tilde{y}_t$ on $\tilde{x}_t$produced the following coefficients (t statistics in parentheses):

$$\tilde{y}_t = \underset{(4.00)}{0.80}\ \tilde{x}_t + e_t. \qquad (1)$$

a.) (10 points) If you used a 5% critical value, would you reject the null hypothesis that the coefficient $\beta = 0$? Would you reject the null hypothesis that $\beta = 1$?

b.) (10 points) A colleague expressed the view that a constant term should be added to the above regression, that is, the colleague wants the researcher to estimate

$$\tilde{y}_t = \beta_0 + \beta_1\tilde{x}_t + \varepsilon_t \qquad (2)$$

Can you calculate numerical values (or if you can't, present mathematical formulas involving $\overline{y}$ and $\overline{x}$) for the estimates of β_0 and β_1 that would result from OLS estimation of (2)? Would you expect the t-statistic for β_1 for (2) to be different from that in (1)? Why?

c.) (20 points) Assume that the vector $(y_t, x_t)'$ is stationary and ergodic with mean $E(y_t, x_t) = (\mu_y, \mu_x)$ and variance-covariance matrix

$$E\left(\begin{bmatrix} E(y_t-\mu_y)^2 & E(y_t-\mu_y)(x_t-\mu_x) \\ E(x_t-\mu_x)(y_t-\mu_y) & E(x_t-\mu_x)^2 \end{bmatrix}\right) = \begin{bmatrix} \sigma_{yy} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{xx} \end{bmatrix}.$$

Calculate the plim of the OLS regression estimate of β_1 in (1), and call this plim π_1. For what economic question (besides being the plim of a regression) does the value of π_1 give the correct answer?

d.) (10 points) Describe in words a statistical test you could use for whether conditional heteroskedasticity might be a problem for regression (1). You do not have to derive any of the properties of the test, just state how you would do it.

e.) (20 points) Describe in words two different methods you might use to deal with conditional heteroskedasticity if the test you performed in (d) suggests that this is a problem. Speculate on how each of these changes might affect the particular numerical estimates reported in (1).

f.) (20 points) Calculate the centered R^2 for regression (1) from the information given. Hint: remember the relation between t statistic and F statistic and alternative ways to calculate an F statistic. Partial credit for just stating the problem correctly.

2.) (25 points total) Another researcher is interested in giving advice to young people about how much more they can expect to earn if they complete another year of schooling based on the estimated value of β in the following OLS regression:

$$y_t = \alpha_1 d_{t1} + \alpha_2 d_{t2} + \alpha_3 d_{t3} + \beta s_t + \varepsilon_t$$

$$\begin{aligned} y_t &= \text{ income of person } t \text{ at age 25} \\ d_{t1} &= 1 \text{ if person } t \text{ is white, 0 otherwise} \\ d_{t2} &= 1 \text{ if person } t \text{ is black, 0 otherwise} \\ d_{t3} &= 1 \text{ if person } t \text{ is race other than black or white, 0 otherwise} \\ s_t &= \text{number of years of schooling obtained by person } t. \end{aligned}$$

a.) (5 points) Would you expect the matrix $\mathbf{X}'\mathbf{X}$ to have full rank in estimating the above equation? Why or why not?

b.) (10 points) Give the single most important reason that would lead you to expect that ε_t is correlated with s_t. Which way would this bias the OLS estimate $\hat{\beta}$?

c.) (10 points) Explain in words one strategy you might use to deal with the problem you raised in part (b).

3.) (135 points total) If $E[\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)] = \mathbf{0}$, 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

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

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

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

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

Suppose you are interested in the value of the $(k \times 1)$ vector $\boldsymbol{\beta}$ in the following regression model:

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

You are concerned about endogeneity, that is, you suspect that not all the elements of $E(\mathbf{z}_t u_t)$ are zero. You have a $(k \times 1)$ vector of instruments $\mathbf{x}_t$ (which may have some elements in common with $\mathbf{z}_t$) for which you believe the following to be true,

$$E(\mathbf{x}_t u_t) = \mathbf{0}$$

$$E(u_t u_{t-v} \mathbf{x}_t \mathbf{x}'_{t-v}) = \begin{cases} \mathbf{\Gamma}_0 & \text{for } v = 0 \\ \mathbf{\Gamma}_1 & \text{for } v = 1 \\ \mathbf{\Gamma}'_1 & \text{for } v = -1 \\ \mathbf{0} & \text{for } v = \pm 2, \pm 3, \ldots \end{cases}$$

where $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_1$ are $(k \times k)$ matrices whose values do not have to be estimated but are known with certainty.

a.) (10 points) Under the conditions stated, would you characterize the instruments $\mathbf{x}_t$ as valid? If not, what other conditions or restrictions would be needed to have valid instruments?

b.) (10 points) Under the conditions stated, would you characterize the instruments $\mathbf{x}_t$ as relevant? If not, what other conditions or restrictions would be needed to have relevant instruments?

c.) (60 points total) Assuming whatever conditions you needed in order to have valid and relevant instruments, write the estimation problem as a special case of GMM. Calculate for this particular case the values of the following.

i.) $\boldsymbol{\theta}$ (5 points)

ii.) $\mathbf{w}_t$ (5 points)

iii.) $\mathbf{h}(\boldsymbol{\theta}, \mathbf{w}_t)$ (5 points)

iv.) $\mathbf{Y}_T$ (5 points)

v.) $\mathbf{g}(\boldsymbol{\theta}, \mathbf{Y}_T)$ (5 points)

vi.) $\hat{\mathbf{D}}'$ (5 points)

vii.) $\mathbf{S}$ (5 points)

viii.) $\hat{\boldsymbol{\theta}}_{GMM}$ (15 points)

ix.) $\mathbf{V}$ (5 points)

x.) $\hat{\mathbf{V}}$ (5 points)

d.) (15 points) Use your answers to part (c) to propose a test of the null hypothesis $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ for $\mathbf{R}$ a known $(m \times k)$ matrix and $\mathbf{r}$ a known $(m \times 1)$ vector. State (but do not derive) the presumed asymptotic distribution of your proposed test statistic.

e.) (40 points total) Suppose you believe that your instruments are weak, even though they are valid and relevant.

i.) (10 points) What does "weak instruments" mean in this context?

ii.) (15 points) What implications would weak instruments have for the estimate of $\boldsymbol{\beta}$ that you proposed in part (c) and for the test you proposed in (d)?

iii.) (15 points) Propose a test you could perform of the null hypothesis $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ that would be robust with respect to weak instruments. You do not have to derive the properties of the test, but state exactly what you would do, give the mathematical formula for the test statistic you would look at, and state the asymptotic distribution you would use to interpret that test statistic.