Econ 220B Final Exam
Winter 2009

DIRECTIONS: No books or notes of any kind are allowed. Answer all questions on separate paper. 250 points are possible on this exam.

1.) (60 points total) The least-squares regression model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ makes a lot of assumptions, including:
(i) $\mathbf{X}'\mathbf{X}$ is of full rank k
(ii) $E(\varepsilon_t^2) = \sigma^2$
(iii) $\text{plim}(\mathbf{X}'\mathbf{X}/T) = \mathbf{Q}$ of full rank k
(iv) $T^{-1/2}\sum_{t=1}^{T}\mathbf{x}_t\varepsilon_t \xrightarrow{L} N(\mathbf{0},\sigma^2\mathbf{Q})$.

In addition, we often make one or more of the following assumptions:
(v) $E(\boldsymbol{\varepsilon}\,|\mathbf{X}) = \mathbf{0}$;
(vi) $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\,|\mathbf{X}) = \sigma^2\mathbf{I}_T$;
(vii) $\boldsymbol{\varepsilon}\,|\mathbf{X} \sim N(\mathbf{0},\sigma^2\mathbf{I}_T)$.
(viii) $\mathbf{x}_t$ is strictly deterministic

Questions (a)-(f) below involve certain statements about properties of the least-squares regression coefficient $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and variance estimate $s^2 = (T-k)^{-1}\mathbf{e}'\mathbf{e}$ for $\mathbf{e} = \mathbf{y} - \mathbf{X}\mathbf{b}$. In each of the questions, you can assume that (i)-(iv) are satisfied. The question is, which further assumptions besides (i)-(iv) would you need in order to prove the stated result? For example, if you think assumptions (i), (ii), and (v) alone are sufficient to prove that $E(\mathbf{b}) = \boldsymbol{\beta}$, then your answer to part (a) should be "v". If you think that (i), (iii), (v) and (vi) are needed to prove that $E(\mathbf{b}) = \boldsymbol{\beta}$, then your answer to part (a) should be "v and vi". Note that you should not give any proofs or arguments– your answer in each case should be some combination of the numerals "v", "vi", "vii", or "viii".
a.) $E(\mathbf{b}) = \boldsymbol{\beta}$
b.) $\mathbf{b}$ is the best linear unbiased estimator of $\boldsymbol{\beta}$
c.) $E[(\mathbf{b}-\boldsymbol{\beta})(\mathbf{b}-\boldsymbol{\beta})'\,|\mathbf{X}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$
d.) $m^{-1}(\mathbf{R}\mathbf{b}-\mathbf{r})'[s^2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\mathbf{b}-\mathbf{r}) \sim F(m, T-k)$ when $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ for $\mathbf{R}$ an $(m \times k)$ matrix
e.) $\mathbf{b} \xrightarrow{p} \boldsymbol{\beta}$
f.) $(\mathbf{R}\mathbf{b}-\mathbf{r})'[s^2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\mathbf{b}-\mathbf{r}) \xrightarrow{L} \chi_m^2$ when $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ for $\mathbf{R}$ an $(m \times k)$ matrix

2.) (50 points total) Consider a regression of a scalar y_t on a constant and another scalar x_t:

$$y_t = \alpha + \gamma x_t + \varepsilon_t.$$

a.) Find an expression for the value of the vector $\hat{\boldsymbol{\beta}} = (\hat{\alpha}, \hat{\gamma})'$ such that the sum of squared residuals

$$\sum_{t=1}^{T} (y_t - \hat{\alpha} - \hat{\gamma} x_t)^2,$$

is as small as possible. (Note: you will receive full credit just for correctly writing down the expression for $\hat{\boldsymbol{\beta}}$, so if you remember it, you do not have to re-derive it here).

b.) Recall the following formulas:

$$\text{sample means:} \quad \bar{x} = T^{-1} \sum_{t=1}^{T} x_t \quad \text{and} \quad \bar{y} = T^{-1} \sum_{t=1}^{T} y_t$$

$$\text{sample variances:} \quad S_{xx} = T^{-1} \sum_{t=1}^{T} (x_t - \bar{x})^2 \quad \text{and} \quad S_{yy} = T^{-1} \sum_{t=1}^{T} (y_t - \bar{y})^2$$

$$\text{sample covariance:} \quad S_{xy} = T^{-1} \sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})$$

$$\text{sample correlation:} \quad r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}.$$

Find an expression for your value for $\hat{\gamma}$ in part (a) in terms of $\bar{x}, \bar{y}, S_{xx}, S_{yy}, S_{xy}$, and r_{xy}. HINT: This calculation and the ones that follow will be easier if you exploit a general property of regressions that include a constant term.

c.) The R^2 from the regression is defined as

$$R^2 = 1 - \frac{\sum_{t=1}^{T} (y_t - \hat{\alpha} - \hat{\gamma} x_t)^2}{\sum_{t=1}^{T} (y_t - \bar{y})^2}.$$

Find an expression for R^2 in terms of $\bar{x}, \bar{y}, S_{xx}, S_{yy}, S_{xy}$, and r_{xy}. Try to simplify as much as you can.

d.) What is the expression for the ordinary OLS F test of the hypothesis $\hat{\gamma} = 0$?

e.) Find the relation between your expression for the F statistic in part (d) and the value of R^2 in part (c).

3.) (40 points total) A given researcher is interested in using a cross-section data set to study the effect of a state's education spending on the earnings of residents of the state. Let s_i denote the average education expenditures (as a fraction of state income) for state i over 1990-1999 and let y_i denote the average income per person in that state over the same period. The researcher proposes to estimate a regression of the form

$$y_i = \beta_1 + \beta_2 s_i + \varepsilon_i.$$

The researcher proposes to use the estimated value of β_2 to advise policy-makers as to how much they might expect to be able to increase earning per person in the state if they were to allocate more resources to education.

a.) Explain in words the possible source of simultaneous equations bias for this example. Would you expect this to result in the OLS estimate $\hat{\beta}_2$ being biased upward or downward?

b.) The researcher was impressed by your answer to part (a) and remembers something from graduate school about how this might be fixed if she had an observation on some other variable x_i, where x_i is a scalar that is referred to as an "instrument" for s_i. She remembers that the instrument x_i is supposed to satisfy satisfy certain properties. *Briefly* explain to the researcher what it would mean for the instrument x_i to be relevant for this example. Give your answer both by writing down a particular equation and in words. Make sure your equation uses the notation from *this* example and not some general result you remember from class or books.

c.) Next please explain to the researcher what is required in order for the instrument x_i to be valid, again using both an equation and words.

d.) Assuming that x_i is both relevant and valid, give an expression for the estimate that the researcher should use to infer the value of the vector $\boldsymbol{\beta} = (\beta_1, \beta_2)'$; (that is, your answer should be an explict formula for your proposed estimate $\hat{\boldsymbol{\beta}}$).

4.) (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

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

where

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

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

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

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

$$\mathbf{D}' = E\left\{ \left. \frac{\partial \mathbf{h}(\boldsymbol{\theta}; \mathbf{w}_t)}{\partial \boldsymbol{\theta}'} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right\}.$$

In this question you are asked to apply these results to maximum likelihood estimation and quasi-MLE of the following model,

$$y_t | x_t, x_{t-1}, ..., x_1, y_{t-1}, y_{t-2}, ..., y_1 \sim N(0, \alpha x_t),$$

for which the log of the conditional density is

$$\log f(y_t | x_t, x_{t-1}, ..., x_1, y_{t-1}, y_{t-2}, ..., y_1) = -(1/2)\log(2\pi) - (1/2)\log(\alpha x_t) - \frac{y_t^2}{2\alpha x_t}$$

and the conditional log likelihood function is

$$\mathcal{L} = -(T/2)\log(2\pi) - (1/2)\sum_{t=1}^{T} \log(\alpha x_t) - \sum_{t=1}^{T} \frac{y_t^2}{2\alpha x_t}.$$

Thus, conditional on the past values of the variable y_t and explanatory variable x_t, the variable y_t is normally distributed with mean zero and a variance that depends linearly on x_t. The explanatory variable x_t is strictly positive and the strictly positive parameter α is going to be estimated by maximum likelihood. In performing the following calculations, you may find it helpful to note that this structure implies $y_t/\sqrt{x_t} \sim N(0, \alpha)$. You may also draw upon the results that if $z \sim N(0, \theta)$, then $E(z^2) = \theta$, $E(z^4) = 3\theta^2$, and $E\{[z^2 - E(z^2)]^2\} = 2\theta^2$.

a.) Show that y_t is a martingale difference sequence.

b.) Calculate the score associated with observation t and show that it has expectation zero when evaluated at the true value α_0.

c.) Find the first-order conditions that characterize the maximum likelihood estimate $\hat{\alpha}$.

d.) Apply the central limit theorem directly to your expression for $\hat{\alpha}$ in part (c) to find the asymptotic distribution of $\hat{\alpha}$.

e.) Use your answers to (b) and (c) to interpret $\hat{\alpha}$ as an example of a GMM estimator.

f.) Show that for this example $\mathbf{S}$ corresponds to the scalar $S = 1/(2\alpha_0^2)$

g.) Find the values of $\mathbf{D}$ and $\mathbf{V}$ for this example to verify that the GMM formula for $\mathbf{V}$ is the same as the asymptotic variance you proposed in part (d).

h.) Suppose instead that you believed that $y_t|x_t, x_{t-1}, ..., x_1, y_{t-1}, y_{t-2}, ..., y_1$ had mean zero, variance αx_t, but not necessarily a Gaussian distribution. Could you still give a quasi-MLE justification for using the estimate $\hat{\alpha}$? Explain.

i.) Describe an alternative estimate of the variance $\hat{\mathbf{V}}$ derived from the general GMM results provided above that would be appropriate for this quasi-MLE interpretation of the estimate.