

QUANTITATIVE METHODS

COMPREHENSIVE EXAMINATION

WINTER 1998

There are three (3) Sections to the examination. Write your answers for each section in a separate bluebook.

All candidates must answer two (2) questions from each section. Each question has the same number of points. There are a total of four (4) hours for the exam.

Section One — Use A Separate Bluebook for this Section.

1. Consider the Gaussian Moving Average:

$$Y_t = \alpha + \theta V_{t-1} + V_t,$$

where $V_t \sim N(0, 1)$.

- (a) Assuming $Y_0 = 0, V_0 = 0$ find the parameters of conditional (i.e., on this value of Y_0) multivariate normal distribution of Y_1, Y_2, Y_3 .
 - (b) Find the exact sampling distribution of the sample mean \bar{Y}_T when $T = 3$, i.e., one observes three values (Y_1, Y_2, Y_3) , **the process is started from its equilibrium distribution at time zero** and $\theta = 0.5$.
 - (c) Find the exact sampling distribution of \bar{Y}_T if $\theta = -1$ and $Y_0 = 0, V_0 = 0$ across all samples and T is finite.
2. Consider the bivariate time series produced by the following Markovian mixture of distributions:

$$Y_{1t} \sim B(3, 2) \text{ if } X_t = 1 \text{ or } Y_{1t} \sim B(2, 3) \text{ if } X_t = -1.$$

and

$$\ln(Y_{2t}) \sim N(-\ln(Y_{1t}), 1) \text{ if } X_t = 1 \text{ or } \ln(Y_{2t}) \sim N(-\ln(Y_{1t}) - 2, 4) \text{ if } X_t = -1.$$

Where

$$P[X_t = 1 | X_{t-1} = 1] = P[X_t = -1 | X_{t-1} = -1] = 0.75.$$

$B(\alpha, \beta)$ is the beta distribution with density:

$$\left(\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \right)^{-1} y_1^{\alpha-1} (1-y_1)^{\beta-1},$$

and moments given by

$$E[Y_1^m] = \frac{\Gamma(m+\alpha)\Gamma(m+\delta)}{\Gamma(\alpha)\Gamma(m+\alpha+\delta)},$$

when $\alpha + m > 0$. Also if $\ln(Y_2) \sim N(\mu, \sigma^2)$ then

$$E[Y_2^m] = \exp(m\mu + 0.5m^2\sigma^2).$$

- (a) Find $E[Y_{it}|X_t = 1]$ and $E[Y_{it}|X_t = -1]$ for $i = 1, 2$.
- (b) Find $E[Y_{1t}Y_{2t}|X_t = 1]$ and $E[Y_{1t}Y_{2t}|X_t = -1]$.
- (c) Find $E[Y_{it}]$ $i = 1, 2$ and $E[Y_{1t}Y_{2t}]$.
- (d) Suppose that at time 0 you are told that $P[X_0 = 1] = 0.5$. Find the posterior belief about X_1 if you are told that $Y_{11} = \exp(2)/\exp(3)$, $\ln(Y_{21}) = 1$.
- (e) Find $E[Y_{12}|Y_{11} = \exp(2)/\exp(3), \ln(Y_{21}) = 1]$ and $E[Y_{22}|Y_{11} = \exp(2)/\exp(3), \ln(Y_{21}) = 1]$.
3. The random variables X and Y are jointly distributed with finite second moments. Let $\epsilon = \ln(Y) - E[\ln(Y)|X]$ and $U = \ln(Y) - E^*[\ln(Y)|X]$, where $E[\exp(Y)|X]$ is the conditional expectation function and $E^*[\ln(Y)|X]$ is the best linear predictor.
- (a) Prove that $E[\epsilon^2] \leq E[U^2]$.
- (b) Suppose that X and Y have a joint normal distribution. Let $f(X)$ be any non-affine function of $\exp(X)$. If $V = \ln(Y) - f(X)$ prove that $E[V^2] > E[U^2]$.
- (c) Is this last result necessarily true if the marginal distributions of $\ln(X)$ and $\ln(Y)$ are normal? If not give a counter-example.

— End of Section One —

Section Two — Use A Separate Bluebook for this Section.

1. Consider the least squares estimator \mathbf{b} for the normal regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

in relation to each of the following two-step estimation procedures.

Each of these procedures begins with k preliminary regressions of each regressor \mathbf{X}_i , $i = 1, \dots, k$ on another set of $n > k$ variables \mathbf{Z} . Represent this set of estimated regressions in matrix form as

$$\mathbf{X} = \mathbf{Z}\hat{\Gamma} + \hat{U}.$$

- (a) Define $\hat{\beta}$ as the least squares estimator from regressing a series \mathbf{y} on the fitted residuals \hat{U} . That is

$$\mathbf{y} = \hat{U}\hat{\beta} + \hat{\varepsilon}.$$

What model misspecification is $\hat{\beta}$ robust to? Can you combine the estimators \mathbf{b} and $\hat{\beta}$ to test for this misspecification? If so, describe the resulting test statistic, and it's (approximate) distribution.

- (b) Suppose that $\mathbf{X}'\mathbf{Z} \approx 0$. How does this affect the estimators \mathbf{b} and $\hat{\beta}$? How does this affect the test that you proposed in (1a)? Can you find another means of testing the underlying misspecification, that is robust to this orthogonality of \mathbf{X} and \mathbf{Z} ?
- (c) Next, let $\tilde{\beta}$ denote the estimator resulting from regressing the series \mathbf{y} on the fitted values from the preliminary regressions, namely $\hat{\mathbf{X}} = \mathbf{Z}\hat{\Gamma}$. What model misspecification is $\tilde{\beta}$ robust to? Can you combine the estimators \mathbf{b} and $\tilde{\beta}$ to test for this misspecification? If so, describe the resulting test statistic, and it's (approximate) distribution under the null hypothesis.
- (d) Again consider the case where $\mathbf{X}'\mathbf{Z} \approx 0$. How does this condition affect $\tilde{\beta}$, its relation to \mathbf{b} , and the test you proposed in (1c)? Can you find another means of testing this second, underlying misspecification, that is robust to this orthogonality of \mathbf{X} and \mathbf{Z} ?
- (e) Consider an alternative procedure, of regressing \mathbf{y} on both $\hat{\mathbf{X}}$ and \hat{U} . Denote the coefficients in this regression as $\hat{\delta}_1$ and $\hat{\delta}_2$ respectively. That is

$$\mathbf{y} = \hat{\mathbf{X}}\hat{\delta}_1 + \hat{U}\hat{\delta}_2 + \hat{\nu}.$$

Which set of coefficients, if either, should be interpreted as estimates of $\boldsymbol{\beta}$? Under what conditions is this interpretation valid? Relate these estimators $\hat{\delta}_1$ and $\hat{\delta}_2$ to those introduced in the preceding parts of this problem.

2. Consider the conditional moment specifications

$$\begin{aligned} E(y_t | \mathbf{X}, \mathbf{Z}) &= \mathbf{X}'_t \boldsymbol{\beta}_1, \\ \text{Var}(y_t | \mathbf{X}, \mathbf{Z}) &= \mathbf{Z}'_t \boldsymbol{\beta}_2. \end{aligned}$$

Here

$$X = \begin{pmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_T \end{pmatrix}, \text{ and } Z = \begin{pmatrix} Z'_1 \\ Z'_2 \\ \vdots \\ Z'_T \end{pmatrix},$$

are $(T \times k_1)$ and $(T \times k_2)$ matrices of observations.

- (a) Provide consistent estimators of β_1 and β_2 .
- (b) Suppose that $x_{j1} = z_{j1} = 1, j = 1, \dots, T$. Let $\beta_1^* = (\beta_{12}, \dots, \beta_{1k})'$, and $\beta_2^* = (\beta_{22}, \dots, \beta_{2k})'$, so that $\beta_1 = (\beta_{11}, \beta_1^*)'$ and $\beta_2 = (\beta_{21}, \beta_2^*)'$. Suppose we want to test the hypothesis that

$$H_2 : \beta_2^* = 0.$$

Describe the construction of a test statistic, and give its asymptotic distribution under the null hypothesis.

- (c) Based on part (b) above, suppose you reject the null H_2 .

Can you use this information to improve your estimate of β_1 ? If so, how are you improving it?

Suppose instead that you accept the null hypothesis H_2 . What does this tell you about the properties of your estimator for β_1 ? Again explain.

- (d) Next you want to test

$$H_1 : \beta_1^* = 0.$$

Construct a test of this hypothesis, based on your results from part (c). Specifically, explain how you construct a test statistic for the case that you accept H_2 , and for the case that you reject H_2 .

Specify the limiting distribution of your test statistics for both cases.

- (e) What is a potential problem with the procedure you described in parts (c) and (d)?

How does this affect your testing results in part (d)?

- (f) Explain why you can, or cannot reverse the procedures above. Can you first test H_1 , and base your tests of H_2 on this outcome?

3. Consider the omitted variables problem in the context of time series data. Assume that the true model is given by

$$y_t = x_t\beta + z_t\delta + \varepsilon_t,$$

where both x_t and z_t are positively serially correlated, univariate time series. You may assume for simplicity that x_t and z_t are mean zero, AR(1) series:

$$x_t = \rho_x x_{t-1} + \eta_{1t},$$

$$z_t = \rho_z z_{t-1} + \eta_{2t},$$

where $|\rho_x|$ and $|\rho_z|$ are less than one, and $\eta_t = (\eta_{1t}, \eta_{2t})$ is bivariate white noise, with covariance matrix

$$E(\eta_t \eta_t') = \Sigma_\eta = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}.$$

Instead of estimating the correctly specified regression above, you omit z_t from the regression.

- (a) Derive the probability limit of the least squares estimator \mathbf{b} of β in the misspecified regression $y_t = \alpha + \beta x_t + u_t$. Under what parameter restrictions does $\mathbf{b} \xrightarrow{p} \beta$?
- (b) Derive the means, variances and covariance of x_t and z_t .
- (c) Let $u_t = z_t \delta + \varepsilon_t$. Derive the variance and autocovariances of u_t . Under what conditions does u_t display positive serial correlation?
- (d) Based upon the results for u_t you have derived above in (3c), comment on the ability of tests for residual serial correlation to offer indirect evidence of the omitted variable z_t .
- (e) Based upon your results regarding the consistency of the least squares estimator for \mathbf{b} , in (3a), comment on the analysis you provided in part (3d).

— End of Section Two —

Section Three — Use A Separate Bluebook for this Section.

1. Consider the following two equation model.

$$y_{1i} = \beta_1 + u_{1i},$$

$$y_{2i} = \beta_2 \cdot x_i + u_{2i}.$$

The disturbances (u_{1i}, u_{2i}) are independent of x_i , and have mean zero.

- (a) Why is least squares estimation of β_2 not efficient in general?
- (b) Discuss SUR estimation of β_2 .
- (c) Discuss GMM estimation of β_2 .
- (d) Test the hypothesis $\beta_1 = 1$ at the 5% level. In this part of the exercise you can use the following information: the observations are independent and identically distributed, N (sample size) is 50, $\sum y_{1i} = 150$, $\sum y_{2i} = 50$, $\sum x_i = 100$, $\sum y_{1i}^2 = 500$, $\sum y_{2i}^2 = 90$, $\sum x_i^2 = 100$, $\sum y_{1i}y_{2i} = 40$, $\sum y_{1i}x_i = 60$ and $\sum y_{2i}x_i = 50$. You can also use

$$\begin{pmatrix} 1 & .2 \\ .2 & 1.3 \end{pmatrix},$$

as an estimate of the covariance matrix of $(u_{1i}, u_{2i})'$.

2. Consider the following model. The probability that a binary variable y_i given $x_i = x$ is equal to 1 is $\exp(\alpha + \beta x)/(1 + \exp(\alpha + \beta x))$.

- (a) Estimate α and its standard error assuming that $\beta = 0$. You can use the information that $N = 100$, $\sum y_i = 50$, $\sum x_i = 60$, $\sum x_i y_i = 35$, and $\sum x_i^2 = 60$.
- (b) Test the hypothesis $\beta = 0$ at the 5% level using the same information about the sample as in (a).
- (c) Estimate α and β , using the fact that x_i is also binary.

3. Consider the following Tobit model.

$$y_i^* = x_i' \beta + u_i,$$

$$y_i = y_i^* \cdot 1\{y_i^* > 0\}.$$

The distribution of u_i given x_i is normal with mean zero and variance σ^2 .

- (a) Show that ols is inconsistent for β .
- (b) Discuss two alternative methods for estimating β that are consistent.

— End of Section Three —