Econometrics

Comprehensive Examination

Spring 1998

There are eight (8) questions in this examination. All candidates must answer <u>any</u> four (4) questions. All questions have equal weight. Please answer each question in a separate bluebook.

Good luck and don't panic!

1. Let the log-likelihood function for an unknown parameter vector θ, given a random sample of observations $(z_1, \ldots z_N)$ be

$$l_N(\theta) = \sum_{i=1}^{N} l(\theta, z_i).$$

(a) Define the maximum simulated likelihood (MSL) estimator for θ, $\hat{\theta}_{MSL}$. Be sure to explain carefully any terms you introduce.

(b) Show that $\hat{\theta}_{MSL}$ is in general inconsistent for a finite number of simulation replications, R.

(c) Derive the asymptotic distribution for $\hat{\theta}_{MSL}$. Discuss it's relationship to the asymptotic distribution of $\hat{\theta}_{MLE}$, the maximum likelihood estimator for θ, and explain intuitively why $\hat{\theta}_{MSL}$ is asymptotically efficient if R increases at rate $\sqrt{N}$.

2. Time-series cross-section data is available on an indicator variable y_{it} and a vector of covariates x_{it}. Two alternative models are proposed for p_{it}, the probability the $y_{it} = 1$.

Model 1: $p_{it} = \alpha_i + x_{it}'\beta$

Model 2: $\log(p_{it}/(1-p_{it})) = \alpha_i + x_{it}'\beta$

(a) Discuss the relative merits of these two models.

(b) Carfeully describe an appropriate estimator (or estimators) for β under the two alternatives, and discuss the properties of these estimators.

(c) Suppose that x_{it} includes lagged values of y_{it} (the indicator variable). Briefly discuss the impact on your answer in part (b).

I. For any $(n \times n)$ matrix A the adjoint A^* is another $(n \times n)$ matrix with elements

$$a_{ij}^* = (-1)^{i+j}|A_{ji}|,$$

where A_{ji} is the $(n-1) \times (n-1)$ matrix that results from deleting the jth row and the ith column of A, and $|\cdot|$ denotes the determinant. For instance, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

then the adjoint is given by

$$A^* = \begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix}' = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}.$$

When A^{-1} exists, it is given by

$$A^{-1} = |A|^{-1}A^*.$$

Similar results hold for $(n \times n)$ lag operator matrices $\mathbf{A}(L)$. Let $\mathbf{A}(L)^*$ denote the adjoint.

A. Under what conditions on the VAR

$$\mathbf{A}(L)\mathbf{y}_t = \mathbf{u}_t$$

will

$$\mathbf{A}(L)^{-1} = |\mathbf{A}(L)|^{-1}\mathbf{A}(L)^*?$$

What do these conditions imply about alternative representations of the system? Hint: the relevant alternative representation is sometimes stated as the Wold Decomposition Theorem.

B. Assuming that the conditions you described in (A) hold, use the adjoint matrix $\mathbf{A}(L)^*$ to construct a VARMA — Vector Autoregressive, Moving Average representation for the VAR above:

$$\mathbf{B}(L)\mathbf{y}_t = \Theta(L)\mathbf{u}_t.$$

Identify $\mathbf{B}(L)$ and $\Theta(L)$. Note the order of the autoregressive and moving average components. What additional feature distinguishes this VARMA representation? Hint: examine the autoregressive component.

C. Let $[A]_{ij}$ denote the element of A corresponding to the ith row and jth column. Denote the covariance matrix of $\mathbf{u}_t$ as Σ. Suppose that the VMA component $\Theta(L)$ you found in (B) is first order:

$$\Theta(L) = \Theta_0 + \Theta_1 L.$$

Let $\mathbf{v}_t = \Theta(L)\mathbf{u}_t$. Derive the autocovariance function for the ith element of $\mathbf{v}_t$.

D. Given your results above, provide a univariate representation for the ith element of $\mathbf{y}_t$.

E. Consider the simple bivariate VAR

$$y_{1t} = \beta y_{2t} + u_{1t}$$
$$y_{2t} = y_{2t-1} + u_{2t}.$$

Under what conditions (if any) does this VAR satisfy the conditions you described in part (A)? What is $\mathbf{A}(L)$ for this system? What is $\mathbf{A}(L)^*$?

F. Can you represent the system in (E) in a VARMA specification, as in part (B)? To answer this question, use the adjoint you derived in part (E), and determine whether the resulting VARMA representation is correct or not. Comment on the necessity and/or sufficiency of the conditions you provided in part (A), for the existence of these VARMA representations.

G. Suppose two series have univariate ARMA(1,1) representations

$$y_{1t} = \rho_1 y_{1t-1} + e_{1t} + \theta_1 e_{1t-1}$$
$$y_{2t} = \rho_2 y_{2t-1} + e_{2t} + \theta_2 e_{2t-1}.$$

Assume as well that $\rho_1 \neq \rho_2$ and $\theta_1 \neq \theta_2$. Could these variables have come from a common VAR? Explain your answer.

II. In "Generalized Method of Moments Specification Testing" Newey discusses asymptotically optimal tests. The solid lines in Figure 1 plot chi-squared distributions with five and fifteen degrees of freedom respectively. Suppose that these are the asymptotic distributions of two test statistics T_5 and T_{15}. The dashed lines represent the distributions of these tests under a specific alternative hypotheses. Figure 2 plots the cumulative distribution functions for these two tests, again under the null and alternative hypotheses.

A. The dashed plots in Figure 1 are simple translations of the (solid) null densities. What distributions do such translations correspond to?

B. For 5% size tests, what are the approximate critical values for each test?

C. At these critical values, what are the approximate rejection probabilities for each test, under the specified alternative? Which test has greater power against this alternative?

D. How much of an additional translation of the alternative distribution of T_5 is necessary to make these tests equally powerful asymptotically, at the 5% size? What conclusions can you draw about the effect of noncentrality parameters and degrees of freedom on the power of tests?

Consider the r population moment conditions

$$E(g(z,\beta_0)) = \int g(z,\beta_0) f(z,c_0)dz = 0,$$

and the analogous sample moments

$$g_T(\beta) = \frac{1}{T}\sum g(z_t,\beta).$$

Newey analyzes moment misspecification in terms of deviations from the proposed distribution $f(z,c_0)$, that is, in terms of $c \neq c_0$. Let

$$h(\beta,c) = \int g(z,\beta) f(z,c)dz.$$

Consider the sequence of alternatives

$$c_T = c_0 + \delta/\sqrt{T}.$$

E. Examine the limiting behavior of $\sqrt{T}h(\beta_0,c_T)$. Be explicit in the assumptions you make.

F. Derive a limiting distribution for $\sqrt{T}g_T(\beta_0)$. Relate the mean of this limiting distribution to your result from part (E). Again be explicit in the assumptions you make.

G. Finally, derive the limiting distribution for $\hat{\beta}$, the GMM estimator, based upon the moments $g_T(\cdot)$ and a weighting matrix W_T. Relate the mean of this limiting distribution to the mean you derived in (F). Hint: first expand $g_T(\hat{\beta})$. Then use the first order condition that the GMM estimator satisfies.

H. Consider a special case, where

$$g_T(\beta) = \begin{pmatrix} U(\beta)'Z \\ U(\beta)'X \end{pmatrix} T^{-1}.$$

Here, $U(\beta) = Y - X\beta$ are residuals from a linear model. What weighting matrices W_{ls} and W_{iv} will yield the least squares and the instrumental variables estimators $\hat{\beta}_{ls}$ and $\hat{\beta}_{iv}$ respectively?

I. Are either of these weighting matrices optimal? Explain.

J. Based upon the weighting matrices in (H), relate the limiting distributions of $\hat{\beta}_{ls}$ and $\hat{\beta}_{iv}$ to the limiting behavior of $\sqrt{T} g_T(\beta_0)$.

K. Use your results from (J) to relate the limiting variance of the difference $\hat{\beta}_{ls} - \hat{\beta}_{iv}$ to the limiting variance of $\sqrt{T} g_T(\beta_0)$.

L. Discuss the relevance of these results to Hausman tests.

Figure 1

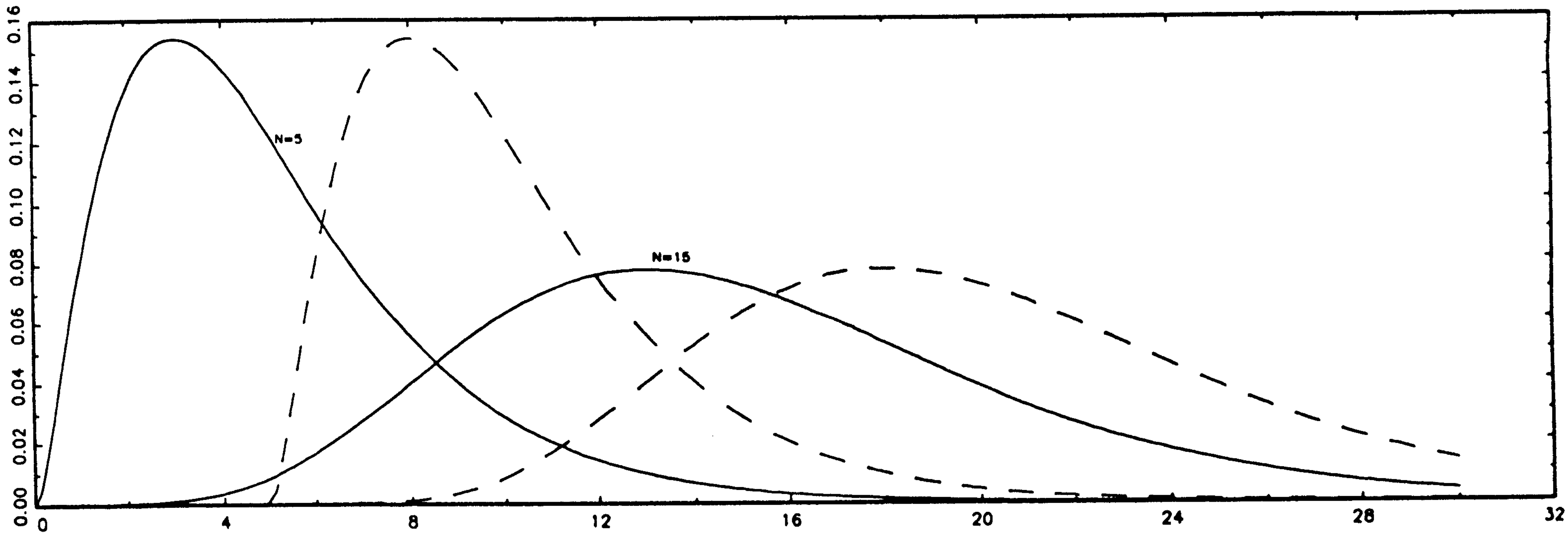

Figure 2