Other variables such as y_{it}^{*} and X_{it}^{*} can be obtained similarly. In this case, $e_T^{\alpha} = Ce_T = (\alpha_1, \alpha_2, \ldots, \alpha_t)'$ where the α_t's can be obtained recursively as follows:

$$\alpha_1 = 1$$

$$\alpha_t = 1 + \sum_{s=1}^{t-1} \lambda_s \alpha_{t-s} \qquad \text{for} \quad t = 2,3,\ldots,q,$$

$$\alpha_t = 1 + \sum_{s=1}^{q} \lambda_s \alpha_{t-s} \qquad \text{for} \quad t = q+1, q+2, \ldots, T. \tag{17}$$

$d^2 = e_T^{\alpha\prime} e_T^{\alpha} = \sum_{t=1}^{T} \alpha_T^2$ and the derivations of Σ^* and $\sigma_\epsilon \Sigma^{*-1/2}$ are the same as before, see (8) through (13). The typical element of y_{it}^{*} can be obtained recursively as in (16), and that of $y^{**} = \sigma_\epsilon \Sigma^{-1/2} y^*$ from (14) with the newly defined α_t's in (17).

REFERENCES

1. Choudhury, A.H. & R.D. St. Louis. A note on Park and Heikes' (1983) modified approximate estimator for the first-order moving-average process. *Journal of Econometrics* 46 (1990): 399–406.
2. Fuller, W.A. & G.E. Battese. Estimation of linear models with cross-error structure. *Journal of Econometrics* 2 (1974): 67–78.
3. Wansbeek, T. & A. Kapteyn. A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics* A11 (1982): 2105–2112.
4. Wansbeek, T. & A. Kapteyn. A note on spectral decomposition and maximum likelihood estimation of ANOVA models with balanced data. *Statistics and Probability Letters* 1 (1983): 213–215.

92.4.4. *Comparison of GLS and OLS for a Linear Regression Model with Noninvertible MA(1) Errors*—Solution,[1] proposed by Luis J. Alvárez and Juan J. Dolado. Let the DGP be

$$y_t = \mu + u_t \qquad (t = 1,2,\ldots,T)$$

$$u_t = e_t - e_{t-1}; \qquad e_0 = 0,\ E(e_t^2) = \sigma^2.$$

(1) In matrix notation the var-cov matrix of $\hat{\mu}$ (OLS estimator) and $\tilde{\mu}$ (GLS estimator) are

$$V(\hat{\mu}) = \sigma^2 (i'\mathrm{i})^{-1} (i'\Omega i)(i'i)^{-1} \tag{1}$$

$$V(\tilde{\mu}) = \sigma^2 (i'\Omega^{-1} i)^{-1}, \tag{2}$$

where $i' = (1,1,\dots,1)$, and

$$\Omega = \begin{bmatrix} 1 & -1 & 0 & 0\dots0 \\ -1 & 2 & -1 & 0\dots0 \\ 0 & -1 & 2 & -1\dots0 \\ & \dots & \dots & \dots \\ 0 & 0 & 0\dots & 2 \end{bmatrix}$$

and

$$\Omega^{-1} = \begin{bmatrix} T & T-1 & T-2 & T-3\dots1 \\ T-1 & T-1 & T-2 & T-3\dots1 \\ T-2 & T-2 & T-2 & T-3\dots1 \\ & \dots & \dots & \dots \\ 1 & 1 & 1 & 1\dots1 \end{bmatrix} = \frac{1}{T-1} \begin{pmatrix} w^{11} & \vdots & w^{12} \\ \cdots & & \cdots \\ w^{21} & \vdots & \Omega^{22} \end{pmatrix}$$

Hence,

$$i'\Omega i = 1 \tag{3}$$

$$i'\Omega^{-1}i = \frac{T(T+1)(2T+1)}{6}. \tag{4}$$

Remark 1. To obtain (4), note that $i'\Omega^{-1}i$ is the sum of the elements of Ω^{-1}. Denote such a sum by S_T and, correspondingly, the sum of the elements of Ω^{22} by S_{T-1}. Then, it follows that

$$S_T = s + S_{T-1}; \qquad S_1 = 1, \tag{5}$$

where s is the sum of w^{11}, w^{12}, and

$$w^{21}\left(= \frac{T(T+1)}{2} + \frac{T(T-1)}{2} = T^2 \right).$$

Hence, from (5), $S_T = \sum_1^T t^2 = T(T+1)(2T+1)/6$. Thus

$$V(\hat{\mu}) = \sigma^2/T^2 \tag{6}$$

$$V(\tilde{\mu}) = \sigma^2\, 6/[T(T+1)(2T+1)], \tag{7}$$

where $V(\hat{\mu}) > V(\tilde{\mu})$ for $T \geq 2$.

Therefore, the limiting distributions of $\hat{\mu}$ and $\tilde{\mu}$ are

$$T(\hat{\mu} - \mu) \Rightarrow N(0,\sigma^2) \tag{8}$$

$$T^{3/2}(\tilde{\mu} - \mu) \Rightarrow N(0,3\sigma^2) \tag{9}$$

by Lyapunov's CLT.

Remark 2. The OLS estimator is $Op(T^{-1})$ ("super-consistent"), while the GLS estimator is $Op(T^{-3/2})$ ("hyper-consistent"). The intuition behind these properties is as follows.

On the one hand, application of OLS yields

$$\hat{\mu} = \mu + \sum_{1}^{T} \frac{(e_t - e_{t-1})}{T} = \mu + \frac{e_T}{T},$$

since e_T is $Op(1)$, then (8) follows.

On the other hand, denote $\tilde{y}_t = y_t/\Delta(= y_t + y_{t-1} + \cdots + y_1)$, then GLS is equivalent to OLS in the model

$$\tilde{y}_t = \mu t + e_t.$$

Hence,

$$\tilde{\mu} = \mu + \sum_{1}^{T} t\epsilon_t \Big/ \sum_{1}^{T} t^2,$$

since $\sum_1^T t\epsilon_t$ is $Op(T^{-3/2})$ and $\sum_1^T t^2$ is $Op(T^{-3})$; then (9) follows.

(2) Theorem 1 in Kruskal [1] says that $\hat{\mu}$ and $\tilde{\mu}$ are the same if $\Omega \times \epsilon\Lambda$ for all $x \in \Lambda$, where x (the regressor) is assumed to lie in the linear manifold Λ. In this case $x = i$, thus

$$\Omega i = \begin{bmatrix} 0 \\ 0 \\ . \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ . \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ . \\ 1 \\ 0 \end{bmatrix} \tag{10}$$

and clearly

$$R(\Omega i) \neq R(i),$$

where the symbol R signifies the range space of a matrix. Thus, it follows by Kruskal's theorem that GLS and OLS are not equivalent.

The graphs in Figure 1 represent the recursive OLS ($\hat{\mu}$) and GLS ($\tilde{\mu}$) estimates up to a sample size of 100 observations.

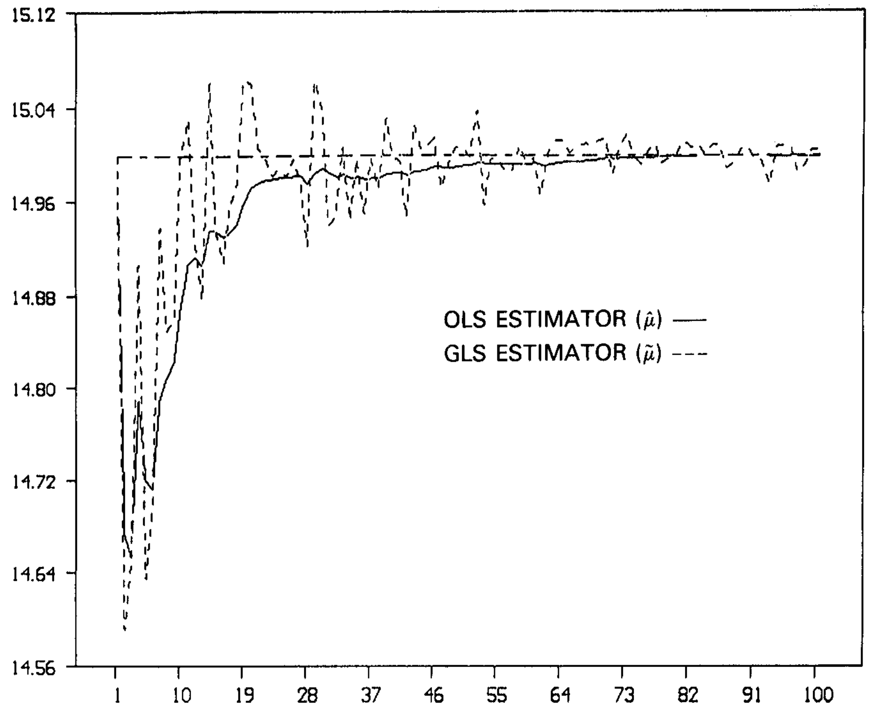

FIGURE 1. DGP: $y_t = 15 + u_t$, $u_t = e_t - e_{t-1}$ $(t = 1, \ldots, T)$, $e_o = 0$, $e_t \sim$ n.i.d. (0,1).

NOTE

1. An excellent solution has been proposed independently by In Choi, the poser of the problem.

REFERENCE

1. Kruskal, W. When are Gauss-Markov and least squares estimators identical? A coordinate free approach. *Annals of Mathematical Statistics* 39 (1968): 70–75.

92.4.5. *Tabulation of Farebrother's Test for Linear Restriction*—Solution,[1] proposed by Jean-Marie Dufour and Sophie Mahseredjian. We consider h separate linear regression models of the form:

$$y_j = X_j\beta_j + \varepsilon_j,\ \varepsilon_j \sim N[0, \sigma_j^2 I_{n_j}], \qquad j = 1, \ldots, h, \tag{1}$$

where y_j is an $n_j \times 1$ vector of observations on a dependent variable, X_j is an $n_j \times k_j$ fixed matrix such that $1 \leq \text{rank}(X_j) = k_j < n_j$, ε_j is an $n_j \times 1$ vec-