a. $\iota' = \phi_4^{1/2}\iota'\Omega^{-1/2}$, and
b. X has an intercept if and only if $\tilde{X} = \Omega^{-1/2}X$ has an intercept, we have

$$\begin{aligned}\iota'e &= \iota'(y - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y),\\ &= \phi_4^{1/2}\iota'\Omega^{-1/2}(y - X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y),\\ &= \phi_4^{1/2}\iota'(I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}')\tilde{y},\\ &= 0,\end{aligned}$$

with $\tilde{y} \equiv \Omega^{-1/2}y$. Since $\iota'e = 0$, $\bar{e} = 0$ and the desired result follows.

NOTE

A very good solution has been proposed by Badi H. Baltagi, the poser of the problem.

REFERENCES

1. Goldberger, A.S. Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association* 57 (1962): 369-375.
2. Wansbeek, T.J. & A. Kapteyn. The separation of individual variation and systematic change in the analysis of panel data. *Annales de l'Insee* 30-31 (1978): 659-680.
3. Wansbeek, T.J. & A. Kapteyn. A note on spectral decomposition and maximum likelihood estimation in ANOVA models with balanced data. *Statistics and Probability Letters* 1 (1983): 213-215.

88.1.2. *Asymptotic Properties of OLS and GLS*—Solution, proposed by Juan J. Dolado, Bank of Spain.

A(i) The OLS($\hat{\beta}$) and GLS($\tilde{\tilde{\beta}}$) estimators are defined by:

$$\hat{\beta} = \sum_t ty_t \Big/ \sum_t t^2 \qquad \text{with} \qquad \text{var}(\hat{\beta}) = \sigma_\epsilon^2 \sum_t \sum_s t\omega_{ts}s \Big/ \left(\sum_t t^2\right), \tag{1}$$

$$\tilde{\tilde{\beta}} = \sum_t \sum_s t\omega^{ts}y_s \Big/ \sum_t \sum_s t\omega^{ts}s \qquad \text{with } \text{var}(\tilde{\tilde{\beta}}) = \sigma_\epsilon^2 \Big/ \sum_t \sum_s t\omega^{ts}s, \tag{2}$$

where $\omega_{ts}(\omega^{ts})$ is the (t,s) element of $\Omega(\Omega^{-1})$, where $\Omega = E(uu')$ is a tridiagonal matrix with $1 + \theta^2$ on the leading diagonal, except the first element which is equal to 1, and θ on the off diagonals. Since in this case var($\hat{\beta}$) and var($\tilde{\tilde{\beta}}$) are not the same, we will say that both estimators are not equivalent in finite samples although, as we will see below, they are asymptotically equivalent.

(ii) To simplify derivations in this and in the next sections, we will use the following lemma, where "$\rightarrow$" denotes weak convergence in distribution.

LEMMA. *If $\alpha(L)$ is a possibly infinite polynomial in the lag operator L, such that $\alpha(z)$ has all of its roots outside the unit circle, then*

$$T^{-3/2}\Sigma t\alpha(L)\epsilon_t \rightarrow N[0,\sigma_\epsilon^2\alpha(1)^2/3].$$

Proof. Under the condition stated for $\alpha(L)$, we can write

$\alpha(L) = \alpha(1) + \alpha^*(L)(1 - L)$,

where $\alpha(1) \neq 0$ and $\alpha^*(1) \neq 0$. Then

$T^{-3/2}\Sigma t\alpha(L)\epsilon_t = T^{-3/2}\Sigma t\alpha(1)\epsilon_t + O(1)$,

since $T^{-1/2}\Sigma\alpha^*(L)\epsilon_t$ has a nondegenerate asymptotic normal distribution by means of an ordinary CLT for weakly dependent variables. Finally, since $E(t\epsilon_t) = 0$ and $\lim \Sigma t^2/T^3 = \frac{1}{3}$, it follows from straight application of a CLT for i.i.d. variables that $T^{-3/2}\Sigma t\epsilon_t \to N[0,\sigma_\epsilon^2/3]$.

Remark: $\sigma_\epsilon^2\alpha(1)^2 = 2\pi f(0)$, where $f(0)$ is the spectrum at frequency zero. ■

We can now derive the asymptotic distributions of $\hat{\beta}$ and $\tilde{\tilde{\beta}}$ by application of the lemma.

Choosing $\alpha(L) = 1 + \theta L$, which verifies the required condition, it follows that

$$T^{3/2}(\hat{\beta} - \beta) = \frac{\Sigma t(1 + \theta L)\epsilon_t}{\Sigma t^2} \cdot \frac{T^3}{T^{3/2}} \to N[0,3\sigma_\epsilon^2(1 + \theta)^2]. \tag{3}$$

Choosing $\alpha(L) = (1 + \theta)^{-1}$ s.t. $\alpha_j = \theta^j(-1)^j$ verifies also the condition, hence

$$T^{3/2}(\tilde{\tilde{\beta}} - \beta) = \frac{T^3}{T^{3/2}}\,\frac{\Sigma t^*\epsilon_t}{\Sigma t^{*2}} = \frac{\Sigma t(1 + \tilde{\tilde{\theta}}L)^{-1}\epsilon_t}{\Sigma[t(1 + \tilde{\tilde{\theta}}L)^{-1}]^2} \cdot \frac{T^3}{T^{3/2}} \to N[0,3\sigma_\epsilon^2(1 + \theta)^2], \tag{4}$$

where $\tilde{\tilde{\theta}}$ is any consistent estimator of θ, for instance that obtained by choosing the stable root from the quadratic $\hat{\rho}_1\theta^2 - \theta + \hat{\rho}_1 = 0$, where $\hat{\rho}_1$ is the first-order autocorrelation of the residuals obtained from applying OLS to the regression equation. Note that we can apply the previous Lemma in (4), when feasible GLS are used, since the difference $(\tilde{\tilde{\theta}} - \theta)$ is $O(T^{-1/2})$, and hence it is asymptotically negligible where multiplied by $T^{3/2}$. The transformed regressor t is obtained recursively from $t^* = t - \tilde{\tilde{\theta}}(t - 1)^*$, $1^* = 1$.

(iii) The results stem from the fact that the spectrum of stationary error term is constant on the elements of the spectrum of the regressors. This condition is clearly satisfied by a trend (Grenander and Rosenblatt [1]), but also by integrated regressors as shown by Phillips and Park [3].

B(iv) The OLS estimator is given by

$$\hat{\beta} = \beta + \frac{\Sigma x_t e_t}{\Sigma x_t^2} = \beta + \frac{\Sigma(\alpha t + v_t)(u_t + \theta u_{t-1})}{\Sigma(\alpha t + v_t)^2} = \beta + O(T^{-3/2}) \tag{5}$$

since $\Sigma t u_t$ is $O(T^{3/2})$, $\Sigma v_t u_t$ is $O(T)$, Σv_t^2 is $O(T)$, and Σt^2 is $O(T^3)$. Hence, $p \lim \hat{\beta} = \beta$. Note that $\hat{\beta}$ is "super-consistent" since the variables

(y_t, x_t) are cotrended, i.e., both contain a linear trend and the linear combination $y_t - \beta x_t$ is stationary.

(v) From (5) and the Lemma in (ii), choosing $\alpha(L) = (1 + \theta L)$, we have

$$T^{3/2}(\hat{\beta} - \beta) = \frac{\alpha \Sigma t u_t (1 + \theta L) + \Sigma V_t u_t (1 + \theta L)}{\Sigma(\alpha t + v_t)^2} \cdot \frac{T^3}{T^{3/2}}$$
$$\to N[0, 3\sigma_u^2((1 + \theta)/\alpha)^2]. \qquad (6)$$

(vi) The model can be written as $y = \beta x_t + u_t(1 + \theta L)$. Hence, choosing $\alpha(L) = (1 + \theta L)^{-1}$,

$$\tilde{\tilde{\beta}} = \beta + \frac{\Sigma x_t^* u_t}{\Sigma x_t^{*2}} = \beta + \frac{\Sigma(\alpha t + v_t)(1 + \theta L)^{-1} u_t}{\Sigma[(\alpha t + v_t)(1 + \theta L)^{-1}]^2} = \beta + O(T^{-3/2}) \qquad (7)$$

by the same arguments as before, where now $x_t^* = x_t - \theta x_{t-1}^*$, $x_1^* = x_1$. Hence, $p \lim \tilde{\tilde{\beta}} = \beta$. Also, since the effect of the lag polynomial in the leading term asymptotically becomes $(1 + \theta)$, it is clear that the distribution of $T^{3/2}(\tilde{\tilde{\beta}} - \beta)$ is that given in (6).

Remark: If the trend is substituted by an $I(1)$ variable, i.e., $\Delta x_t = v_t$, then the asymptotic distribution is non-normal unless $\sigma_{uv} = 0$, and $(\hat{\beta} - \beta)$ and $(\tilde{\tilde{\beta}} - \beta)$ are $O(T^{-1})$. Hence, simultaneity, unlike in the trend case, affects the asymptotic distribution.

Editor's comment: An excellent solution has been proposed by Peter C.B. Phillips, the poser of the problem. In particular, his proof of A(i), which is reproduced below, makes use of Kruskal's theorem [2].

Let $x' = (1, 2, \ldots, T)$, and

$$V = \operatorname{var}(u) = \begin{bmatrix} 1 & \theta & 0 & \cdots & 0 \\ \theta & 1+\theta^2 & \theta & \cdots & 0 \\ 0 & \theta & 1+\theta^2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1+\theta^2 \end{bmatrix}.$$

Then

$$Vx = \begin{bmatrix} 1 + 2\theta \\ 2 + 4\theta + 2\theta^2 \\ 3 + 6\theta + 3\theta^2 \\ \vdots \\ T + 2(T-1)\theta + (T-1)\theta^2 \\ T + (T-1)\theta + T\theta^2 \end{bmatrix} = x(1+\theta)^2 + \begin{bmatrix} -\theta^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ -(T+1)\theta \end{bmatrix}$$

and clearly

$$R(Vx) \neq R(x),$$

where the symbol R signifies the range space of a matrix. It follows by Kruskal's theorem [2] that OLS and GLS are not equivalent on (1) when $\theta \neq 0$. This answers (i).

However, Vx is close to being collinear with x, and as $T \to \infty$ the effects of the differences diminish since they affect only two observations.

REFERENCES

1. Grenander, U. & M. Rosenblatt. *Statistical analysis of stationary time series*. New York: Wiley, 1957.
2. Kruskal, W. When are Gauss-Markov and least-squares estimators identical? A coordinate free approach. *Annals of Mathematical Statistics* 39 (1968): 70–75.
3. Phillips, P.C.B. & J.Y. Park. Asymptotic properties of OLS and GLS in regressions with integrated regressors. *Journal of the American Statistical Society* 83 (1988): 111–115.