94.3.4. *Spurious Regression in Forecast-Encompassing Tests*—Solution, proposed by Peter C.B. Phillips. From regression (2), we have

$$\hat{b} = \sum_1^n y_t t \bigg/ \sum_1^n t^2 = n^{-1/2}\left(n^{-5/2}\sum_1^n y_t t\right)\bigg/\left(n^{-3}\sum_1^n t^2\right)$$

and then the usual I(1) regression asymptotics give us the limit

$$n^{1/2}\hat{b} \to_d \int_0^1 B(r) r\, dr \bigg/ \int_0^1 r^2\, dr = 3\int_0^1 B(r) r\, dr = \xi_b, \quad \text{say,}$$

where $B(r) \equiv BM(\omega^2)$ and $\omega^2 = \sigma_\varepsilon^2(\sum_{j=0}^\infty c_j)^2$ is the long-run variance of u_t.

Observe that $\tilde{y}_{n+h} = y_n$ for all $h = 1, \dots, N$, so that $y_{n+h} - \tilde{y}_{n+h} = \sum_{n+1}^{n+h} u_t = x_h$, say, which is an integrated process. Because $\hat{y}_{n+h} = \hat{b}(n+h)$, the "forecast-encompassing" regression is

$$x_h = \gamma^* \hat{b}(n+h) + \text{error}, \qquad h = 1, \dots, N,$$

and therefore

$$\gamma^* = \hat{b}\sum_{h=1}^N x_h(n+h)/\hat{b}^2 \sum_{h=1}^N (n+h)^2 = \sum_1^N x_h(n+h)/\hat{b}\sum_1^N (n+h)^2.$$

Note that

$$N^{-5/2}\sum_1^N x_h(n+h) = (n/N)N^{-3/2}\sum_1^N x_h + N^{-5/2}\sum_1^N x_h h$$

$$\to_d (1/\tau)\int_0^1 B_f(s)\, ds + \int_0^1 B_f(s) s\, ds,$$

where $B_f(r) \equiv BM(\omega^2)$ and is independent of $B(r)$, a fact that is easily established by using the BN decomposition (see Phillips and Solo, 1992), and

$$N^{-3}\sum_{1}^{N}(n+h)^2 = n^2/N^2 + nN(N+1)/N^3 + (N+1)(2N+1)/6N^2$$

$$\to (1/\tau)^2 + (1/\tau) + \tfrac{1}{3} = (3+3\tau+\tau^2)/3\tau^2.$$

Hence,

$$\gamma^* = \frac{\left\{N^{-5/2}\sum_{1}^{N} x_h(n+h)\right\} n^{1/2}N^{-1/2}}{(n^{1/2}\hat{b})\left(N^{-3}\sum_{1}^{N}(n+h)^2\right)}$$

$$\to_d \frac{\left\{(1/\tau)\int_0^1 B_f(s)\,ds + \int_0^1 B_f(s)s\,ds\right\}(1/\tau)^{1/2}}{\xi_b(3+3\tau+\tau^2)/3\tau^2} = \xi_\gamma, \quad \text{say.}$$

Next, the t-ratio for γ^* is

$$t_\gamma = \gamma^*/s_{\gamma^*}, \qquad s_{\gamma^*}^2 = s^2\left[\sum_{1}^{N}(\hat{b}(n+h))^2\right]^{-1},$$

with $s^2 = (1/N)\sum_1^N\{x_h - \gamma^*\hat{b}(n+h)\}^2$. Now,

$$N^{-1}s^2 = N^{-2}\sum_{1}^{N}x_h^2 - \gamma^{*2}N^{-2}\sum_{1}^{N}(\hat{b}(n+h))^2$$

$$= N^{-2}\sum_{1}^{N}x_h^2 - \gamma^{*2}(n^{1/2}\hat{b})^2(N/n)N^{-3}\sum_{1}^{N}(n+h)^2$$

$$\to_d \int_0^1 B_f^2 - \xi_\gamma^2\xi_b^2\tau(3+3\tau+\tau^2)/3\tau^2.$$

Thus,

$$t_\gamma = \frac{\gamma^*\left[|\hat{b}|\left\{\sum_{1}^{N}(n+h)^2\right\}^{1/2}\right]}{N^{1/2}(N^{-1/2}s)}$$

$$= \frac{\gamma^* n^{-1/2}(n^{1/2}|b|)N^{3/2}\left\{N^{-3}\sum_{1}^{N}(n+h)^2\right\}^{1/2}}{N^{1/2}(N^{-1/2}s)}$$

$$= O_p(1)n^{-1/2}N = O_p(n^{1/2}),$$

and so t_γ diverges at the rate $O_p(n^{1/2})$ as $n \to \infty$.

It follows that a t-ratio test for the significance of γ^* in the "forecast-encompassing" regression

$$y_{n+h} - \tilde{y}_{n+h} = \gamma^*\hat{y}_{n+h} + \text{error}$$

always leads to a rejection of the null that $\gamma = 0$ as $n \to \infty$. Thus, the forecast $\hat{y}_{n+h}$ from the spurious regression of y_t on the linear trend always turns out to be significant in explaining the forecast error $y_{n+h} - \tilde{y}_{n+h}$ that is obtained from the true model that has a unit root. In this sense, the forecast-encompassing regression is a spurious one.

REFERENCE

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.