93.4.5. *Nonlinear Testing and Forecasting Asymptotics with Potential Rank Failure*—Solution, proposed by Peter C.B. Phillips.

(a) The maximum likelihood estimator $\hat{\theta}$ is obtained by applying nonlinear least squares to equation (1), i.e., by minimizing the sum of squares so that

$$\sum_{t=1}^{T} (y_t - \alpha x_{1t} - \beta x_{2t} - \alpha\beta x_{3t})^2 = \sum_{t=1}^{T} (y_t - x_t'\gamma(\theta))^2, \text{ say.}$$

Under the given assumptions, we have the limit theory

$$\sqrt{T}(\hat{\theta} - \theta) \to_d N(0, V_\theta),$$

where

$$V_\theta = \sigma^2 \left[\frac{\partial\gamma(\theta)'}{\partial\theta} M \frac{\partial\gamma(\theta)}{\partial\theta'} \right]^{-1} \quad \text{and} \quad M = \lim_{T\to\infty} (T^{-1}X'X) = I_3.$$

Because

$$\frac{\partial\gamma(\theta)'}{\partial\theta} = \begin{bmatrix} 1 & 0 & \beta \\ 0 & 1 & \alpha \end{bmatrix},$$

we obtain

$$V_\theta = \sigma^2\begin{bmatrix} 1+\beta^2 & \alpha\beta \\ \alpha\beta & 1+\alpha^2 \end{bmatrix}^{-1} = \frac{\sigma^2}{1+\alpha^2+\beta^2}\begin{bmatrix} 1+\alpha^2 & -\alpha\beta \\ -\alpha\beta & 1+\beta^2 \end{bmatrix}.$$

(b) Let $\Psi(\theta) = \alpha\beta$. Then the Wald statistic for testing

$H_0\colon\Psi(\theta) = 0$

is

$$W_T = T\Psi(\hat\theta)'\left[\frac{\partial\gamma(\hat\theta)}{\partial\theta'}\,\hat V_\theta\,\frac{\partial\Psi(\hat\theta)'}{\partial\theta}\right]^{-1}\Psi(\hat\theta),$$

where

$$\Psi(\hat\theta) = \hat\alpha\hat\beta,$$

$$\hat V_\theta = \hat\sigma^2\left[\frac{\partial\gamma(\hat\theta)'}{\partial\theta}\,M_{xx}\,\frac{\partial\gamma(\hat\theta)}{\partial\theta'}\right]^{-1},$$

$$\frac{\partial\Psi(\hat\theta)}{\partial\theta'} = (\hat\beta,\hat\alpha),$$

and $\hat\sigma^2 = T^{-1}\Sigma_1^T(y_t - x_t'\gamma(\hat\theta))^2$ is the MLE of the error variance σ^2 in the model.

Because $M_{xx} \to I_3$ as $T \to \infty$, the limit distribution of W_T is the same as that of the statistic where M_{xx} is replaced by I_3, namely

$$\hat W_T = \frac{T(\hat\alpha\hat\beta)^2(1+\hat\alpha^2+\hat\beta^2)}{\hat\sigma^2(\hat\alpha^2+\hat\beta^2)}. \tag{*}$$

Now, under H_0 we have $\alpha\beta = 0$ and

$$\sqrt{T}\hat\alpha\hat\beta \to_d N\left(0,\frac{\partial\Psi(\theta)}{\partial\theta'}\,V_\theta\,\frac{\partial\Psi(\theta)'}{\partial\theta}\right) \equiv N(0,\sigma^2(\alpha^2+\beta^2)/(1+\alpha^2+\beta^2)). \tag{**}$$

Because $(\hat\alpha,\hat\beta) \to_p (\alpha,\beta)$, it follows from (*) and (**) that

$$W_T\hat W_T \to_d \chi_1^2,$$

except when $\alpha = \beta = 0$. Note that in this case (where $\alpha\beta = 0$ and H_0 is satisfied), $\hat\theta' = (\hat\alpha,\hat\beta) \to_p (0,0)$ and

$$\sqrt{T}\hat\theta \to_d N(0,\sigma^2 I_2). \tag{†}$$

Thus, W_T, $\hat{W}_T$ have the same limit distribution as

$$\frac{T(\hat{\alpha}\hat{\beta})^2}{\sigma^2(\hat{\alpha}^2 + \hat{\beta}^2)} = \frac{\{(T^{1/2}\hat{\alpha}/\hat{\sigma})(T^{1/2}\hat{\beta}/\hat{\sigma})\}^2}{\{(\hat{\alpha}/\hat{\sigma})^2 + (\hat{\beta}/\hat{\sigma})^2\}} \to_d \chi_\alpha^2\chi_\beta^2/(\chi_\alpha^2 + \chi_\beta^2), \qquad (\dagger\dagger)$$

where χ_α^2 and χ_β^2 are both χ_1^2 (chi-square with one degree of freedom) variates and are statistically independent in view of (†).

(c) The forecast $\hat{y}_{T+1}$ is

$$\hat{y}_{T+1} = \hat{\alpha}x_{1T+1} + \hat{\beta}x_{2T+1} + \hat{\alpha}\hat{\beta}x_{3T+1},$$

and forecast error is

$$y_{T+1} - \hat{y}_{T+1} = u_{T+1} + (\alpha - \hat{\alpha})x_{1T+1} + (\beta - \hat{\beta})x_{2T+1} + (\alpha\beta - \hat{\alpha}\hat{\beta})x_{3T+1},$$

which is asymptotically equivalent to

$$u_{T+1} + x'_{T+1}(\partial\gamma(\theta)/\partial\theta')(\hat{\theta} - \theta) = u_{T+1} + \bar{x}'_{T+1}(\hat{\theta} - \theta), \text{ say.}$$

The asymptotic variance of the forecast error is then

$$\hat{\text{var}}(y_{T+1} - \hat{y}_{T+1}) = \hat{\sigma}^2 + \{1 + \hat{\bar{x}}'_{T+1}(\hat{\bar{X}}'\hat{\bar{X}})^{-1}\hat{\bar{x}}_{T+1}\}, \qquad (\dagger\dagger\dagger)$$

where

$$\hat{\bar{x}}_{T+1} = x'_{T+1}\partial\gamma(\hat{\theta})/\partial\theta' = (x_{1T+1} + \hat{\beta}x_{3T+1}, x_{2T+1} + \hat{\alpha}x_{3T+1}),$$

and

$$\hat{\bar{X}}'\hat{\bar{X}} = \Sigma_1^T \hat{\bar{x}}_t\hat{\bar{x}}'_t = \frac{\partial\gamma(\hat{\theta})'}{\partial\theta}(X'X)\frac{\partial\gamma(\hat{\theta})}{\partial\theta'}.$$

(d) When $\theta = 0$ (i.e., $\alpha = \beta = 0$), the model is simply $y_t = u_t \equiv$ i.i.d. $N(0, \sigma^2 I)$. Note that the null hypothesis

$$H_0 : \alpha\beta = 0$$

holds in this case. We now have

$$\sqrt{T}(\hat{\theta} - \theta) = \sqrt{T}\hat{\theta} \to_d N(0, \sigma^2 I_2),$$

$$W_T \to_d \chi_\alpha^2\chi_\beta^2/(\chi_\alpha^2 + \chi_\beta^2)$$

as in (††), and

$$\hat{\text{var}}(y_{T+1} - \hat{y}_{T+1}) = \hat{\sigma}^2\{1 + \hat{\bar{x}}'_{T+1}(\hat{\bar{X}}'\hat{\bar{X}})^{-1}\hat{\bar{x}}_{T+1}\},$$

as in (†††).