$$e_2 = (e_{12}, e_{22}, e_{32}, \ldots, e_{n2})$$
$$e_3 = (e_{13}, e_{23}, e_{33}, \ldots, e_{n3})$$
$$\ldots$$
$$e_T = (e_{1T}, e_{2T}, e_{3T}, \ldots, e_{nT})$$

Thus, the first bootstrap sample of the residuals might be

$$e_1^* = (e_{13}, e_{23}, e_{33}, \ldots, e_{n3})$$
$$e_2^* = (e_{15}, e_{25}, e_{35}, \ldots, e_{n5})$$
$$\ldots$$
$$e_T^* = (e_{12}, e_{22}, e_{32}, \ldots, e_{n2})$$

The point is that you resample in such a way as to maintain the contemporaneous relationships among the regression residuals. As in Example 2, construct a bootstrap series using the resampled values of $\{e_t^*\}$. Once you obtain the average value of the t-statistics for the first bootstrap sample, repeat the entire process several thousand times.

Efron and Tibshirani (1993) presented an extremely accessible treatment of bootstrapping. You may also download a programming manual that illustrates some bootstrapping techniques (at no charge) from my Web site or from the Wiley Web site for this text.

APPENDIX 4.2: DETERMINATION OF THE DETERMINISTIC REGRESSORS

Sometimes the appropriate null and alternative hypotheses are unclear. As indicated in the text, you do not want to lose power by including a superfluous deterministic regressor in a unit root test. However, omitting a regressor that is actually in the data-generating process leads to a misspecification error. Fortunately, Sims, Stock, and Watson (1990) provided a second rule that is helpful in selecting the appropriate set of regressors:

> **Rule 2:** If the data-generating process contains any deterministic regressors (*i.e.*, an intercept or a time trend) and the estimating equation contains these deterministic regressors, inference on all coefficients can be conducted using a t-test or an F-test. This is because a test involving a single restriction across parameters with different rates of convergence is dominated asymptotically by the parameters with the slowest rates of convergence.

While the proof is beyond the scope of this text, the point is that the nonstandard Dickey–Fuller distributions are needed only when you include deterministic regressors not in the actual data-generating process. Hence, if the DGP is known to contain the