These are often called **augmented Dickey-Fuller tests**, or **ADF tests**. They were proposed originally by Dickey and Fuller (1979) under the assumption that the error terms follow an AR process of known order. Subsequent work by Said and Dickey (1984) and Phillips and Perron (1988) showed that they are asymptotically valid under much less restrictive assumptions. Consider the test regressions (20.05), (20.06), (20.07), or (20.11). We can write any of these regressions as

$$\Delta y_t = \boldsymbol{X}_t\boldsymbol{\beta} + (\alpha - 1)y_{t-1} + u_t, \tag{20.14}$$

where $\boldsymbol{X}_t$ consists of whatever set of nonstochastic regressors is included in the test regression: nothing at all for (20.06), a constant for (20.07), a constant and a linear trend for (20.05), and so on.

Now suppose, for simplicity, that the error term u_t in (20.14) follows the stationary AR(1) process $u_t = \rho u_{t-1} + \varepsilon_t$. Then (20.14) would become

$$\begin{aligned}\Delta y_t &= \boldsymbol{X}_t\boldsymbol{\beta} - \rho\boldsymbol{X}_{t-1}\boldsymbol{\beta} + (\rho + \alpha - 1)y_{t-1} - \alpha\rho y_{t-2} + \varepsilon_t \\ &= \boldsymbol{X}_t\boldsymbol{\beta}^* + (\rho + \alpha - 1 - \alpha\rho)y_{t-1} + \alpha\rho(y_{t-1} - y_{t-2}) + \varepsilon_t \qquad (20.15) \\ &= \boldsymbol{X}_t\boldsymbol{\beta}^* + (\alpha - 1)(1 - \rho)y_{t-1} + \alpha\rho\Delta y_{t-1} + \varepsilon_t. \qquad (20.16)\end{aligned}$$

We are able to replace $\boldsymbol{X}_t\boldsymbol{\beta} - \rho\boldsymbol{X}_{t-1}\boldsymbol{\beta}$ by $\boldsymbol{X}_t\boldsymbol{\beta}^*$ in (20.15), for some choice of $\boldsymbol{\beta}^*$, because every column of $\boldsymbol{X}_{t-1}$ lies in $\mathcal{S}(\boldsymbol{X})$. This is a consequence of the fact that $\boldsymbol{X}_t$ can include only such deterministic variables as a constant, a linear trend, and so on (see Section 10.9). Thus each element of $\boldsymbol{\beta}^*$ is a linear combination of the elements of $\boldsymbol{\beta}$.

Equation (20.16) is a linear regression of Δy_t on $\boldsymbol{X}_t$, y_{t-1}, and Δy_{t-1}. This is just the original regression (20.14), with one additional regressor, Δy_{t-1}. Adding this regressor has caused the serially dependent error term u_t to be replaced by the serially independent error term ε_t. The ADF version of the τ statistic, which we will refer to as the τ' statistic, is simply the ordinary t statistic for the coefficient on y_{t-1} in (20.16) to be zero. If the serial correlation in the error terms of (20.14) were fully accounted for by an AR(1) process, this τ' statistic would have exactly the same asymptotic distribution as the ordinary DF τ statistic for the same specification of $\boldsymbol{X}_t$. The fact that the coefficient on y_{t-1} is $(\alpha-1)(1-\rho)$ rather than $\alpha-1$ does not matter. Because it is assumed that $|\rho| < 1$, this coefficient can be zero only if $\alpha = 1$. Thus a test for the coefficient on y_{t-1} to be zero is equivalent to a test for $\alpha = 1$.

It is evidently very easy to compute τ' statistics using regressions like (20.16), but it is not so easy to compute the corresponding z' statistics. If the coefficient of y_{t-1} were multiplied by n, the result would be $n(\hat{\alpha} - 1)(1 - \hat{\rho})$ rather than $n(\hat{\alpha} - 1)$. This test statistic clearly would not have the same asymptotic distribution as z. Thus, in order to compute a valid z' statistic from regression (20.16), it is necessary to divide the coefficient of y_{t-1} by $1-\hat{\rho}$; see Dickey, Bell, and Miller (1986).