than the other may be seen as a deficiency of these tests. That is so only if one misinterprets their nature. Nonnested hypothesis tests are specification tests, and since there is almost never any reason a priori to believe that either of the models actually generated the data, it is appropriate that nonnested tests, like other model specification tests, may well tell us that neither model seems to be compatible with the data.

It is important to stress that the purpose of nonnested tests is *not* to choose one out of a fixed set of models as the "best" one. That is the subject of an entirely different strand of the econometric literature, which deals with criteria for **model selection**. We will not discuss the rather large literature on model selection in this book. Two useful surveys are Amemiya (1980) and Leamer (1983), and an interesting recent paper is Pollak and Wales (1991).

It is of interest to examine more closely the case in which both models are linear, that is, $\boldsymbol{x}(\boldsymbol{\beta}) = \boldsymbol{X}\boldsymbol{\beta}$ and $\boldsymbol{z}(\boldsymbol{\gamma}) = \boldsymbol{Z}\boldsymbol{\gamma}$. This will allow us to see why the J and P tests (which in this case are identical) are asymptotically valid and also to see why these tests may not always perform well in finite samples. The J-test regression for testing H_1 against H_2 is

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \alpha \boldsymbol{P}_Z \boldsymbol{y} + \text{residuals}, \tag{11.16}$$

where $\boldsymbol{P}_Z = \boldsymbol{Z}(\boldsymbol{Z}^\top\boldsymbol{Z})^{-1}\boldsymbol{Z}^\top$ and $\boldsymbol{b} = (1-\alpha)\boldsymbol{\beta}$. Using the FWL Theorem, we see that the estimate of α from (11.16) will be the same as the estimate from the regression

$$\boldsymbol{M}_X\boldsymbol{y} = \alpha \boldsymbol{M}_X\boldsymbol{P}_Z\boldsymbol{y} + \text{residuals}. \tag{11.17}$$

Thus, if $\acute{s}$ denotes the OLS estimate of σ from (11.16), the t statistic for $\alpha = 0$ will be

$$\frac{\boldsymbol{y}^\top\boldsymbol{P}_Z\boldsymbol{M}_X\boldsymbol{y}}{\acute{s}(\boldsymbol{y}^\top\boldsymbol{P}_Z\boldsymbol{M}_X\boldsymbol{P}_Z\boldsymbol{y})^{1/2}}. \tag{11.18}$$

First of all, notice that when only one column of $\boldsymbol{Z}$, say $\boldsymbol{Z}_1$, does not belong to $\mathcal{S}(\boldsymbol{X})$, it must be the case that

$$\mathcal{S}(\boldsymbol{X}, \boldsymbol{P}_Z\boldsymbol{y}) = \mathcal{S}(\boldsymbol{X}, \boldsymbol{Z}) = \mathcal{S}(\boldsymbol{X}, \boldsymbol{Z}_1).$$

Therefore, the J-test regression (11.16) must yield exactly the same SSR as the regression

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \delta\boldsymbol{Z}_1 + \text{residuals}. \tag{11.19}$$

Thus, in this special case, the J test is equal in absolute value to the t statistic on the estimate of δ from (11.19).

When two or more columns of $\boldsymbol{Z}$ do not belong to $\mathcal{S}(\boldsymbol{X})$, this special result is no longer available. If the data were actually generated by H_1, we can replace $\boldsymbol{y}$ in the numerator of (11.18) by $\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{u}$. Since $\boldsymbol{M}_X\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}$, that numerator becomes

$$\boldsymbol{\beta}^\top\boldsymbol{X}^\top\boldsymbol{P}_Z\boldsymbol{M}_X\boldsymbol{u} + \boldsymbol{u}^\top\boldsymbol{P}_Z\boldsymbol{M}_X\boldsymbol{u}. \tag{11.20}$$