the test based on (10.98) is testing against a less general alternative than the usual form of the test. When $x_t(\boldsymbol{\beta})$ is linear, (10.97) can be written as

$$(1-\rho L)y_t = \boldsymbol{X}_t\boldsymbol{\beta} - \delta\boldsymbol{X}_{t-1}\boldsymbol{\beta} + \varepsilon_t, \tag{10.99}$$

which is in general (but not when $l = 1$) more restrictive than equation (10.89). Thus consideration of the nonlinear regression case reveals that there are really two different tests of common factor restrictions when the original model is linear. The first, which tests (10.88) against (10.89), is the F test (10.92). It will have l degrees of freedom, where $1 \leq l \leq k$. The second, which tests (10.88) against (10.99), is the t test of $d = 0$ in the Gauss-Newton regression (10.98). It will always have one degree of freedom. Either test might perform better than the other, depending on how the data were actually generated; see Chapter 12. When $l = 1$, the two tests will coincide, a fact that it may be a good exercise to demonstrate.

10.10 Instrumental Variables and Serial Correlation

So far in this chapter, we have assumed that the regression function $\boldsymbol{x}(\boldsymbol{\beta})$ depends only on exogenous and predetermined variables. However, there is no reason for serially correlated errors not to occur in models for which current endogenous variables appear in the regression function. As we discussed in Chapter 7, the technique of instrumental variables (IV) estimation is commonly used to obtain consistent estimates for such models. In this section, we briefly discuss how IV methods can be used to estimate univariate regression models with errors that are serially correlated and to test for serial correlation in such models.

Suppose that we wish to estimate the model (10.12) by instrumental variables. Then, as we saw in Section 7.6, the IV estimates may be obtained by minimizing, with respect to $\boldsymbol{\beta}$ and ρ, the criterion function

$$\big(\boldsymbol{y} - \boldsymbol{x}'(\boldsymbol{\beta},\rho)\big)^{\top}\boldsymbol{P}_W\big(\boldsymbol{y} - \boldsymbol{x}'(\boldsymbol{\beta},\rho)\big), \tag{10.100}$$

where the regression function $\boldsymbol{x}'(\boldsymbol{\beta},\rho)$ is defined by (10.13), and $\boldsymbol{P}_W$ is the matrix that projects orthogonally onto the space spanned by $\boldsymbol{W}$, a suitable matrix of instruments. The IV form of the Gauss-Newton regression can be used as the basis for an algorithm to minimize (10.100). Given suitable regularity conditions on $x_t(\boldsymbol{\beta})$, and assuming that $|\rho| < 1$, these estimates will be consistent and asymptotically normal. See Sargan (1959) for a full treatment of the case in which $\boldsymbol{x}(\boldsymbol{\beta})$ is linear.

The only potential difficulty with this IV procedure is that one has to find a "suitable" matrix of instruments $\boldsymbol{W}$. For asymptotic efficiency, one always wants the instruments to include all the exogenous and predetermined variables that appear in the regression function. From (10.13), we see that more such variables appear in the regression function $x_t'(\boldsymbol{\beta}, \rho)$ for the transformed model than in the original regression function $x_t(\boldsymbol{\beta})$. Thus the optimal choice of instruments may differ according to whether one takes account of serial correlation or assumes that it is absent.

To make this point more clearly, let us assume that the original model is linear, with regression function

$$x_t(\boldsymbol{\beta}) = \boldsymbol{Z}_t\boldsymbol{\beta}_1 + \boldsymbol{Y}_t\boldsymbol{\beta}_2, \tag{10.101}$$

where $\boldsymbol{Z}_t$ is a row vector of explanatory variables that are exogenous or predetermined, and $\boldsymbol{Y}_t$ is a row vector of current endogenous variables; the dimension of $\boldsymbol{\beta} \equiv [\boldsymbol{\beta}_1 \vdots \boldsymbol{\beta}_2]$ is k. The regression function for the transformed model is then

$$x_t'(\boldsymbol{\beta}, \rho) = \rho y_{t-1} + \boldsymbol{Z}_t\boldsymbol{\beta}_1 + \boldsymbol{Y}_t\boldsymbol{\beta}_2 - \rho\boldsymbol{Z}_{t-1}\boldsymbol{\beta}_1 - \rho\boldsymbol{Y}_{t-1}\boldsymbol{\beta}_2. \tag{10.102}$$

In (10.101), the only exogenous or predetermined variables were the variables in $\boldsymbol{Z}_t$. In (10.102), however, they are y_{t-1} and the variables in $\boldsymbol{Z}_t$, $\boldsymbol{Z}_{t-1}$, and $\boldsymbol{Y}_{t-1}$ (the same variables may occur in more than one of these, of course; see the discussion of common factor restrictions in the previous section). All these variables would normally be included in the matrix of instruments $\boldsymbol{W}$. Since the number of these variables is almost certain to be greater than $k+1$, it would not normally be necessary to include any additional instruments to ensure that all parameters are identified.

For more discussion of the estimation of single linear equations with serially correlated errors and current endogenous regressors, see Sargan (1959, 1961), Amemiya (1966), Fair (1970), Dhrymes, Berner, and Cummins (1974), Hatanaka (1976), and Bowden and Turkington (1984).

Testing for serial correlation in models estimated by IV is straightforward if one uses a variant of the Gauss-Newton regression. In Section 7.7, we discussed the GNR (7.38), in which the regressand and regressors are evaluated at the restricted estimates, and showed how it can be used to calculate test statistics. Testing for serial correlation is simply an application of this procedure. Suppose we want to test a nonlinear regression model for AR(1) errors. The alternative model is given by (10.12), for observations 2 through n, with the null hypothesis being that $\rho = 0$. In this case, the GNR (7.38) is

$$\tilde{\boldsymbol{u}} = \boldsymbol{P}_W\tilde{\boldsymbol{X}}\boldsymbol{b} + r\boldsymbol{P}_W\tilde{\boldsymbol{u}}_{-1} + \text{residuals}, \tag{10.103}$$

where $\tilde{\boldsymbol{\beta}}$ denotes the IV estimates under the null hypothesis of no serial correlation, $\tilde{\boldsymbol{u}}$ denotes $\boldsymbol{y} - \boldsymbol{x}(\tilde{\boldsymbol{\beta}})$, and $\tilde{\boldsymbol{X}}$ denotes $\boldsymbol{X}(\tilde{\boldsymbol{\beta}})$. This is clearly the IV analog of regression (10.76); if the two occurrences of $\boldsymbol{P}_W$ were removed, (10.76) and (10.103) would be identical. The t statistic on the estimate of r from this regression will be a valid test statistic. This will be true both when (10.103) is estimated explicitly by OLS and when $\tilde{\boldsymbol{u}}$ is regressed on $\tilde{\boldsymbol{X}}$ and $\tilde{\boldsymbol{u}}_{-1}$ using