LAGGED DEPENDENT VARIABLES AND AUTOREGRESSIVE DISTURBANCES

Models with Lagged-Dependent Variables

The reactions of economic agents, such as consumers or investors, to changes in their environment resulting, for example, from changes in prices or incomes, are never instantaneous. The changes are likely to be distributed over time; and positions of equilibrium, if they are ever attained, are likely to be approached gradually.

The slowness to respond may be due to two factors. In the first place, there will be time delays in the transmission and the reception of the information upon which the agents base their actions. In the second place, costs will be entailed in the process of adapting to the new circumstances; and these costs are liable to be positively related to the speed and to the extent of the adjustments. For these reasons, it is appropriate to make some provision in econometric equations for dynamic responses which are distributed over time.

The easiest way of setting an econometric equation in motion is to introduce an element of feedback. This is done by including one or more lagged values of the dependent variable on the right-hand side of the equation to stand in the company of the other explanatory variables. It transpires that, if the current disturbance is unrelated to the lagged dependent variables, then the standard results concerning the consistency of the ordinary least-squares regression procedure retain their validity. This is despite the fact that we can no longer assert that the ordinary least-square estimates of the parameters are unbiased in finite samples.

If the current disturbances and the lagged-dependent variables which are included on the RHS of a dynamic regression equation are not unrelated, then the resulting parameter estimates are liable to suffer from considerable biases. The biases are worst when the variance of the disturbance process is large relative to the variances of the explanatory variables.

The essential nature of the problem can be illustrated via a simple model which includes only a lagged dependent variable and which has no other explanatory variables. Imagine that the disturbances follow a first-order autoregressive process. Then there are two equations to be considered. The first of these is the regression equation

$$y(t) = y(t-1)\beta + \eta(t), \qquad \text{where} \quad |\beta| < 1, \tag{1}$$

and the second is the equation

$$\eta(t) = \rho\eta(t-1) + \varepsilon(t), \qquad \text{where} \quad |\rho| < 1, \tag{2}$$

which describes the autoregressive disturbance process. Here $\varepsilon(t)$ stands for an unobservable white-noise process which generates a sequence of independently and identically distributed random variables which are assumed to be independent of the elements of $y(t)$ which precede them in time. The conditions on the parameters β and ρ are necessary to ensure the stability of the model. That is to say, they are necessary conditions for the attainment of a long-run equilibrium in the dynamic response.

Equations (1) and (2), it will be observed, have the same mathematical form. Using the lag operator L, we may rewrite them, in slightly different forms, as

$$(I - \beta L)y(t) = \eta(t) \quad \text{and} \quad \eta(t) = \frac{\varepsilon(t)}{I - \rho L}. \tag{3}$$

Combining the latter gives

$$(I - \rho L)(I - \beta L)y(t) = \big\{I - (\rho + \beta)L + \rho\beta L^2\big\}y(t) = \varepsilon(t). \tag{4}$$

This is a special case of the equation

$$(I - \beta_1 L - \beta_2 L^2)y(t) = \varepsilon(t) \tag{5}$$

which relates to the regression of the sequence $y(t)$ on itself lagged by one and by two periods. The only restriction which is entailed by writing the equation in the form of (4) derives from the implication that ρ and β are real-valued coefficients. In the case of equation (5), the corresponding values λ_1 and λ_2, which would be obtained by factorising the the polynomial

$$1 + \beta_1 z + \beta_2 z^2 = (1 - \lambda_1 z)(1 - \lambda_2 z), \tag{6}$$

might be complex numbers. In that case, the two equations (4) and (5) would have different implications regarding their dynamic responses to the disturbances in $\varepsilon(t)$.

Now consider the effect of fitting a model with a single lagged value from the sequence $y(t)$ in the role of the explanatory variable. This can be described as the endeavour to estimate the parameter β of equation (1) by applying ordinary least-squares regression to the equation whilst overlooking the serially correlated nature of the disturbance sequence $\eta(t)$.

Both $y(t)$ and $\eta(t)$ are serially correlated sequences which are linked to each other via equation (1). Therefore the current elements of $\eta(t)$ will be correlated with both past, current and future values of $y(t)$. This means that the essential condition on which the consistency of the ordinary least-squares estimator depends is violated.

On substituting the expression $y_t = (\rho + \beta)y_{t-1} - \rho\beta y_{t-2} + \varepsilon_t$ from (4) into the regression formula, we derive the following expression for the estimate:

$$\begin{aligned}\hat{\beta} &= \frac{\sum y_{t-1}y_t}{\sum y_{t-1}^2} \\ &= (\rho + \beta)\frac{\sum y_{t-1}^2}{\sum y_{t-1}^2} - \rho\beta\frac{\sum y_{t-1}y_{t-2}}{\sum y_{t-1}^2} + \frac{\sum y_{t-1}\varepsilon_t}{\sum y_{t-1}^2}.\end{aligned} \tag{7}$$

It is straightforward to take limits in the expression as the sample size T increases. Let $\hat{\beta} \to \delta$ as $T \to \infty$. Then the equation above becomes the equation

$$\delta = (\beta + \rho) - \beta\rho\delta. \tag{8}$$

The final term on the RHS of (7) vanishes since, according to the assumptions, the elements of $\varepsilon(t)$ are uncorrelated with elements of $y(t)$ which precede them in time. Rearranging equation (8) gives the result that

$$\delta = \frac{\rho + \beta}{1 + \rho\beta}. \tag{9}$$

Notice that the expression for δ is symmetric with respect of ρ and β. However, we have tended to regard β as the regression parameter and ρ as the parameter of an autoregressive disturbance process. This distinction now appears to be false. However, if $y(t-1)$ on the RHS of equation (1) were standing in the company of another explanatory variable, say $x(t)$, then the distinction would be a valid one.

Now let us imagine, for the sake of argument, that $\rho \to 0$. Then it is clear that $\delta \to \beta$. Since the variance of the process $\eta(t)$ is related positively to the value of ρ, it can be said that the bias in β is directly related to the variance of the serially-correlated disturbance process. Exactly the same result obtains when $y(t-1)$ is accompanied in the regression equation by other explanatory variables.