LECTURE / DISCUSSION

Time Series Models, Part 2

Introduction

Dynamic models frequently appear in models of time series. Such models explain the current behavior of the dependent variable with contemporaneous and past values of variables. For example, the model

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + u_t$$

specifies that y_t is explained by the contemporaneous value of x , x_t , and its ***lagged*** value, x_{t-1} . It is convenient to consider two types: those models with lagged values of the explanatory variables and those with lagged values of the dependent variable. The former family of models fit into the regression framework that we have been studying: adding x_{t-1} to the list of explanatory variables is analogous to adding a variable like x_t^2 . Models with lagged dependent variables as explanatory variables are special because they introduce a ***random*** variable to the list of explanatory variables.

Such models explain the current behavior of the dependent variable with its own past. For example,

$$y_t = \gamma y_{t-1} + \beta_1 + \beta_2 x_{t2} + \ldots + \beta_K x_{tK} + u_t$$

Some direct ways in which explanatory variables include ***lagged values*** of the dependent variable include models of ***adaptive expectations*** and ***partial adjustment.***

Adaptive Expectations Model

Economic agents routinely make decisions about the future based upon their current expectations about unknown conditions. For example, a household decides whether to retrofit its heating system based in part on the expected prices of fuels and equipment. Thus, in the simplest model we could specify,

$$y_t = \beta_1 + \beta_2 x_t^e + u \quad ,$$

were x_t^e is the expected price. Since expectations about the future are formed on the basis of past behavior, we might model expected x_t^e as

$$x_t^e = x_{t-1}^e + \lambda(x_{t-1} - x_{t-1}^e) \quad , \quad 0 \le \lambda \le 1$$

so that the current expectation is a revision of the previous expectation modified by the observed error in the previous expectation. The dynamics in the formation of expectations (which are not observed by researchers) lead to dynamics in the observed data:

$$y_t = (1 - \lambda)y_{t-1} + \lambda\beta_1 + \lambda\beta_2 x_{t-1} + \varepsilon_t \quad ,$$

where $\varepsilon_t = u_t + (1 - \lambda)u_{t-1}$. This regression function has the simple dynamic form suggested above.

Partial Adjustment Model

A similar phenomenon occurs in a simple model of partial adjustment, in which economic agents cannot adjust fully to changing conditions. For example, as a household ages, the size of the desired residence changes but the costs of adjusting size inhibit actually changing residence size to the preferred level. We could model ***desired*** residence size with a linear regression

$$y_t^* = \beta_1 + \beta_2 x_t + u_t$$

and relate the (unobserved) desired level to the observed level by a partial adjustment process

$$y_t - y_{t-1} = \lambda(y_t^* - y_{t-1}) + \nu_t \ , \quad 0 \leq \lambda \leq 1 \ .$$

This leads to the observable relationship between y and x,

$$y_t = (1 - \lambda)y_{t-1} + \lambda\beta_1 + \lambda\beta_2 x_t + \varepsilon_t \quad ,$$

where $\varepsilon_t = \nu_t + \lambda u_t$.

Distributed Lag Specifications

In general, the dynamic specification of a regression with lagged values of the explanatory variables is not specified clearly by theory. Thus, one is tempted to write

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + \beta_4 x_{t-2} + \ldots + u_t$$

so that the number of lagged terms is infinite. But such specifications are not workable. Economic time series often exhibit near multicollinearity in $x_t, x_{t-1}, x_{t-2}, \ldots$, so that the β's (or w's) cannot be estimated accurately. Econometricians have invented several specifications that are parsimonious. For notational convenience, let

$$y_t = \beta_1 + \beta_2(w_0 x_t + w_1 x_{t-1} + w_2 x_{t-2} + \ldots) + u_t \quad .$$

Geometric Distributed Lag

The geometric distributed lag is the simplest and most parsimonious such specification:

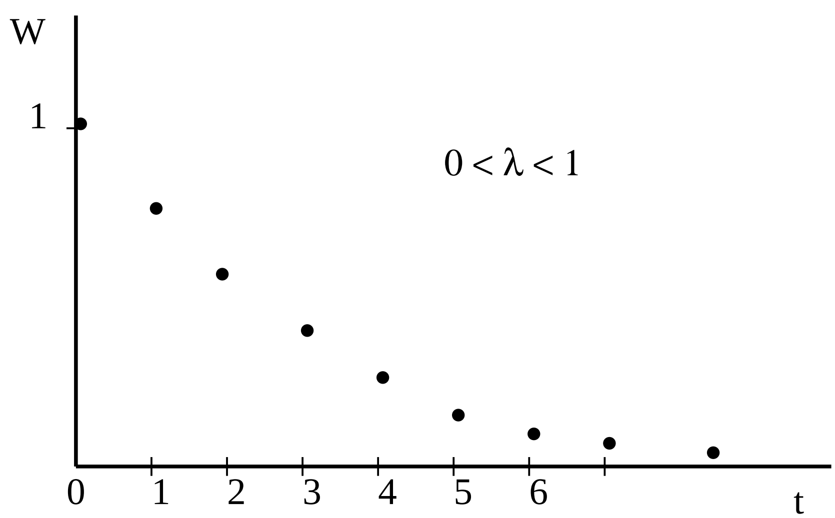
$$w_i = \lambda^i , \quad |\lambda| < 1 \ .$$

To estimate such a model, we rewrite the regression as

$$y_t = \lambda y_{t-1} + \beta_1(1 - \lambda) + \beta_2 x_t + \varepsilon_t \quad ,$$

where $\varepsilon_t = u_t - \lambda u_{t-1}$, and apply an appropriate regression method.

In effect, the list of highly collinear explanatory variables $\{x_{t-1}, x_{t-2},\ldots\}$ has been replaced by the single y_{t-1}, reducing collinearity significantly.



Polynomial Distributed Lag

Another specification writes the distributed lag weights in terms of an underlying polynomial. For example,

$$w_i = c_0 + c_1 i + c_2 i^2 + c_3 i^3 \quad , \quad i = 0,1,2,\ldots,8$$

specifies an eight-term distributed lag that depends on four parameters. The distributed lag coefficients will follow a smooth polynomial shape.

In TSP,

```
OLSQ y c x(4,9,FAR)
```

where 4 = degree of the polynomial + 1; 9 = number of lags + 1; and FAR = keyword for zero restrictions (i.e., you are restricting the last w to be zero). You can also restrict the first w to be zero (NEAR), or do BOTH or NONE.

You must adjust your sample statement (SMPL) to accommodate the lagged periods (at least the number of lags + 1), and the SMPL statement should be inserted just before you use the PDL term (i.e., don’t put it before the LOAD statement).

Lagged Dependent Variables

Two special issues arise when the lagged explanatory variables include the dependent variable: (1) the dynamic stability of the specification and (2) autocorrelation in the disturbances.

Dynamic Stability

Consider the simple autoregressive model,

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \quad ,$$

which can also be written as

$$\begin{aligned} y_t &= \beta_0 + \beta_1\beta_0 + \beta_1^2 y_{t-2} + u_t + \beta_1 u_{t-1} \\ &= \beta_0(1+\beta_1+\ldots+\beta_1^n) + \beta_1^{n-1} y_{t-n} + u_t + \beta_1 u_{t-1} + \ldots + \beta_1^n u_{t-n} \end{aligned}$$

If $|\beta_1| > 1$, then y_t must grow explosively. If $|\beta_1| < 1$, then the process "settles down."

$$\beta_0(1 + \beta_1 + \ldots + \beta_1^n) \rightarrow \frac{\beta_0}{1-\beta_1} \quad ,$$

$$\beta_1^{n-1} y_{t-n} \rightarrow 0 \quad ,$$

and

$$u_t + \beta_1 u_{t-1} + \ldots + \beta_1^n u_{t-n} \rightarrow \mathbb{N}[0, \sigma^2/(1-\beta_1^2)] \quad .$$

A simple way to check for stability in more complicated models is to simulate them. If a dynamic model is unstable, then the simulations will have an explosive pattern. Otherwise, the simulations will converge.

Autocorrelated Disturbances

When there is a lagged dependent explanatory variable and the disturbances are autocorrelated, then OLS regression fails to produce unbiased estimators. This failure is caused by correlation between an explanatory variable, the lagged dependent variable, and the disturbance term. In general,

$$\mathrm{Cov}(u_t,u_{t-1}) \neq 0 \Rightarrow \mathrm{Cov}(u_t,y_{t-1}) \neq 0$$

because u_{t-1} is the disturbance term in y_{t-1}. We will discuss a general approach to overcoming such correlation between explanatory variables and the disturbance term in the lecture on instrumental variables. The quasi first-differencing method that we used earlier can be applied to these models when the disturbances are modeled as autoregressive.

Testing for first-order autoregressive disturbances is slightly different when there is a lagged dependent explanatory variable. Instead of simply regressing OLS fitted residuals on their lagged value, one regresses the residuals on their lagged value ***and*** all of the explanatory variables in the original regression.

```
? basic regression
olsq y y(-1) c x;
olsq @res @res(-1) y(-1) c x;
? check t-stat on @res(-1)
```