Università di Pavia

Dynamic Regression Models

Eduardo Rossi

- $\mathbf{y}_t$ denote an $(n \times 1)$ vector of economic variables generated at time t.
- The collection $\{\mathbf{y}_t, -\infty < t < \infty\}$ is called a (vector-valued) random sequence.
- An economic data set is a finite segment, $\{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_T\}$, of this infinite sequence.
- **Data Generating Process (DGP)**: Joint probability law under which the sequence is generated, embodying all these influences.
- Assume that the data are continuously distributed.

The DGP is completely represented by the conditional density

$$D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right), \quad \mathcal{Y}_{t-1} = \sigma\left(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots\right)$$

$\mathcal{Y}_{t-1}$ is the smallest $\sigma-$ field of events with respect to which the random variables $\mathbf{y}_{t-j}$ are measurable for all $j \geq 0$.

- $D_t(\cdot)$ is allowed to depend on time, because the data are not assumed to be stationary and in particular allowance must be made for features such as seasonal variations, and changes in technology, regulatory regime, ecc.

A dynamic econometric model is a family of functions of the data, of relatively simple form, devised by an investigator, which are intended to mimic aspects of the DGP.
Model is a family of functions

$$\{M(\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{d}_t; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\}, \quad \Theta \subseteq \mathbb{R}^p$$

$$M_D = \text{Model of the complete DGP}$$

$$M_E = \text{Model of } E(\mathbf{y}_t | \mathcal{Y}_{t-1})$$

$$M_V = \text{Model of } Var(\mathbf{y}_t | \mathcal{Y}_{t-1})$$

Models depend on a finite collection of parameters $\boldsymbol{\theta}$ $(p \times 1)$, Θ is the parameter space.

- DGP is not represented as depending on Θ.
- Parameterization is always a feature of the model, not explicitly of the DGP.
- Many different parameterization of DGP are possible.
- **The axiom of correct specification is the assumption that there exists a model element that is identical to the corresponding function of the DGP.**

M_D is correctly specified if there exists $\boldsymbol{\theta}_0 \in \Theta$ such that

$$M_D\left(\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{d}_t; \boldsymbol{\theta}_0\right) = D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right)$$

Similarly M_E and M_V are correct if the conditions

$$M_E\left(\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{d}_t; \boldsymbol{\theta}_0\right) = \int \mathbf{y} D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right) d\mathbf{y}$$

$$\begin{aligned} M_V\left(\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots, \mathbf{d}_t; \boldsymbol{\theta}_0\right) &= \int \mathbf{y}\mathbf{y}' D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right) d\mathbf{y} \\ &\quad - \int \mathbf{y} D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right) d\mathbf{y} \int \mathbf{y}' D_t\left(\mathbf{y}_t \left| \mathcal{Y}_{t-1}\right.\right) d\mathbf{y} \end{aligned}$$

The axiom of correct specification is implausible.

Misspecification is common in practical modelling.

Search for adequate approximation: i.e. M_D, M_E, M_V are true in essentials. Specification tests that allow us to check whether a model fulfils the criteria set.

The dependence of $D_t(\mathbf{y}_t | \mathcal{Y}_{t-1})$ on time can be through its stochastic arguments and/or through the variations in parameters: $\boldsymbol{\theta}_t$. Define

$$\boldsymbol{\theta}_t = f(\boldsymbol{\theta}, \mathbf{d}_t)$$

The variables $\mathbf{d}_t$ are determined outside the economic system.

Dummy variables: Constructed by the investigator to represent a particular source of variation, rather than measured directly.

- Trend functions: t^k, $k \neq 0$;
- Seasonal dummies;
- Intervention dummies: 1 over certain periods, and 0 otherwise, representing particular policy regimes.

The application of dynamic modelling concepts to regression models.

Let y_t be a variable to be modelled and let $\mathbf{z}_t$ $(k \times 1)$ be a vector of explanatory variables, assumed weakly exogenous with respect to the parameters of interest.

The conditioning set for this problem is

$$\mathcal{F}_t = \sigma\left(\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, y_{t-1}, y_{t-2}, \ldots\right)$$

the list of eligible of conditioning variables generally extends beyond those actually playing a role in the model.

Assuming linearity, the models of interest take the general form

$$\begin{aligned} y_t &= \delta' \mathbf{d}_t + \beta_0' \mathbf{z}_t + \beta_1' \mathbf{z}_{t-1} + \ldots + \beta_m' \mathbf{z}_{t-m} \\ &\quad + \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} \\ &\quad + u_t + \theta_1 u_{t-1} + \ldots + \theta_q u_{t-q} \end{aligned}$$

where $E\left(u_t \left| \mathcal{F}_t \right.\right) = 0$. $\mathbf{d}_t$ denotes dummy variables (intercept, seasonals, etc.).
This is an **AutoRegressive Moving Average Distributed Lag** (ARMADL model).

The lagged innovations belong to the conditioning set $\mathcal{F}_t$.
Using the lag operator

$$\alpha(L) y_t = \boldsymbol{\delta}' \mathbf{d}_t + \boldsymbol{\beta}(L)' \mathbf{z}_t + \theta(L) u_t \tag{1}$$

where

$$\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$$

$$\theta(L) = 1 + \theta L + \ldots + \theta_q L^q$$

$$\beta(L) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 L + \boldsymbol{\beta}_2 L^2 + \ldots + \boldsymbol{\beta}_m L^m$$

Explicitly omitting the MA component $\theta(L)$ yields

$$\alpha(L)\, y_t = \delta' \mathbf{d}_t + \boldsymbol{\beta}(L)' \, \mathbf{z}_t + u_t$$

called **AutoRegressive Distributed Lag Model**, which can be estimated by OLS. This is probably the commonest type of model fitted in practice, just because of its simplicity.

The dynamics of a linear time series process can always be expressed in terms of the level and a lag polynomial in the differences. Given a polynomial $\alpha(z)$ of order p, there exists the equivalent representation

$$\alpha(z) = \alpha(1) + \alpha^*(z)(1-z)$$

$$\alpha(1) = \alpha_0 + \alpha_1 + \ldots + \alpha_p$$

where the coefficients of the $p-1$-order polynomial $\alpha^*(z)$ are

$$\alpha_j^* = -\sum_{k=j+1}^{p} \alpha_k \quad j = 0, 1, \ldots, p-1$$

The first and second order cases are

$$\alpha_0 + \alpha_1 z = (\alpha_0 + \alpha_1) - \alpha_1 (1 - z)$$

$$\alpha_0 + \alpha_1 z + \alpha_2 z^2 = (\alpha_0 + \alpha_1 + \alpha_2) - (\alpha_1 + \alpha_2)(1 - z) - \alpha_2 z (1 - z)$$

This is the *Beveridge-Nelson Decomposition*. A further rearrangement yields

$$\alpha(z) = \alpha(1) z + \alpha^{**}(z)(1 - z)$$

where

$$\alpha^{**}(z) = \alpha^{*}(z) + \alpha(1)$$

Transforming $\beta(z)$ similarly, the ARDL model can be written in the so-called Error Correction Model (ECM) form,

$$\alpha^{**}(z)\,\Delta y_t = \boldsymbol{\delta}'\mathbf{d}_t + \boldsymbol{\beta}^{**}(L)'\,\Delta\mathbf{z}_t - \alpha\left(y_{t-1} - \boldsymbol{\theta}'\mathbf{z}_{t-1}\right) + u_t$$

where $\alpha = \alpha(1)$ and $\boldsymbol{\theta} = \boldsymbol{\beta}(1)/\alpha$.

This is only a reparameterization.

The parameters $\boldsymbol{\theta}$ can be thought of as representing the long-run equilibrium relations of the model, those that would prevail if $\mathbf{z}_t$ were constant and $u_t = 0$ for an indefinitely long period.

ARDL(1,1):

$$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + u_t \qquad |\alpha_1| < 1$$

$$y_t = \frac{\beta_0 + \beta_1 L}{1 - \alpha_1 L} z_t + \frac{1}{1 - \alpha_1 L} u_t$$

$$y_t = \beta_0 \sum_{j=0}^{\infty} \alpha_1^j z_{t-j} + \beta_1 \sum_{j=0}^{\infty} \alpha_1^j z_{t-1-j} + \sum_{j=0}^{\infty} \alpha_1^j u_{t-j}$$

$$y_t = \delta(L) z_t + \frac{1}{1 - \alpha_1 L} u_t$$

$$\delta(L) = \frac{\beta_0 + \beta_1 L}{1 - \alpha_1 L}$$

$$(1 - \alpha_1 L)\delta(L) = \beta_0 + \beta_1 L$$

$$
\begin{aligned}
\delta_0 &= \beta_0 \\
\delta_1 &= (\alpha_1 \beta_0 + \beta_1) \\
\delta_j &= \alpha_1 \delta_{j-1} \;\; j \geq 2
\end{aligned}
$$

Static equilibrium (all changes have ceased).

We are treating (y_t, z_t) as jointly stationary. The long-run values are given by the unconditional expectations: $E(y_t)$

$$
\begin{aligned}
y^* &= E(y_t) \\
z^* &= E(z_t).
\end{aligned}
$$

Since $E(u_t) = 0$

$$y^* = \alpha_1 y^* + \beta_0 z^* + \beta_1 z^*$$

$$y^* = \frac{(\beta_0 + \beta_1)}{(1 - \alpha_1)} z^* = \left(\sum_{i=0}^{\infty} \delta_i \right) z^* \equiv k z^*$$

k long-run multiplier (total multiplier) of y with respect to z.

Passing from z^* to $z^* + 1$, the new solution is

$$\delta(1)(z^* + 1) = y^* + \left(\sum_{i=0}^{\infty} \delta_i \right)$$

Impact multiplier

$$\delta_0 = \beta_0$$

instantaneous effect of an increase in z on y.

Interim multipliers

$$\delta_J = \sum_{i=0}^{J} \delta_i \qquad J = 0, 1, 2, \ldots$$

$$\delta_J = \frac{\beta_0(1-\alpha_1) + (\alpha_1\beta_0 + \beta_1)(1-\alpha_1^J)}{1-\alpha_1} \qquad J = 0, 1, 2, \ldots$$

the effect of a unit change in z after J periods.

Standardized Interim Multiplier

$$\delta_J^+ = \frac{\delta_J}{\delta(1)}$$

$$\delta_J^+ = \frac{\beta_0(1-\alpha_1) + (\alpha_1\beta_0 + \beta_1)(1-\alpha_1^J)}{\beta_0 + \beta_1}$$

portion of total adjustment that take place in the first J periods, $J = 0, 1, \ldots$.

When $\delta(1) \neq 0$ and

$$0 \leq \delta_i \leq 1 \qquad i = 1, 2, \ldots$$

The average lag is defined by

$$\mu = \frac{\sum_{j=0}^{\infty} j\delta_j}{\sum_{j=0}^{\infty} \delta_j}$$

ARDL(1,1):

$$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + u_t \qquad |\alpha_1| < 1$$

Nested models:

Static regression	$\alpha_1 = \beta_1 = 0$	$y_t = \beta_0 z_t + u_t$
AR(1)	$\beta_0 = \beta_1 = 0$	$y_t = \alpha_1 y_{t-1} + u_t$
Leading indicator	$\alpha_1 = \beta_0 = 0$	$y_t = \beta_1 z_{t-1} + u_t$
Partial adjustment	$\beta_1 = 0$	$y_t = \alpha_1 y_{t-1} + \beta_0 z_t + u_t$
Distributed lags	$\alpha_1 = 0$	$y_t = \beta_0 z_t + \beta_1 z_{t-1} + u_t$
Common Factors	$\beta_1 = -\beta_0 \alpha_1$	$y_t = +\beta_0 z_t + \epsilon_t \quad \epsilon_t = \alpha_1 \epsilon_{t-1} + u_t$
First Difference	$\alpha_1 = 1 \quad \beta_1 + \beta_0 = 0$	$\Delta y_t = \beta_0 \Delta z_t + u_t$
Dead Start	$\beta_0 = 0$	$y_t = \alpha_1 y_{t-1} + \beta_1 z_{t-1} + u_t$
ECM with homogeneity constraint	$\beta_0 + \beta_1 + \alpha_1 = 1$	$\Delta y_t = \beta_0 \Delta z_t + (\alpha_1 - 1)(y_{t-1} - z_{t-1}) + u_t$

Imposing the restriction

$$\alpha_1 + \beta_0 + \beta_1 = 1$$

$$1 - \alpha_1 = \beta_0 + \beta_1$$

$$y_t \boxed{-y_{t-1}} = \alpha_1 y_{t-1} \boxed{-y_{t-1}} + \beta_0 z_t \boxed{-\beta_0 z_{t-1} + \beta_0 z_{t-1}} + \beta_1 z_{t-1} + u_t$$

$$\Delta y_t = -y_{t-1} + \alpha_1 y_{t-1} + (\beta_0 z_t - \beta_0 z_{t-1}) + (\beta_0 z_{t-1} + \beta_1 z_{t-1}) + u_t$$

$$\Delta y_t = -(1 - \alpha_1) y_{t-1} + \beta_0 \Delta z_t + (\beta_0 + \beta_1) z_{t-1} + u_t$$

$$\Delta y_t = -(1 - \alpha_1) y_{t-1} + \beta_0 \Delta z_t + (1 - \alpha_1) z_{t-1} + u_t$$

$$\Delta y_t = \beta_0 \Delta z_t + (1 - \alpha_1)(z_{t-1} - y_{t-1}) + u_t$$

In the ECM formulation, parameters describing the extent of short-run adjustment to disequilibrium are immediately provided by the regression.

ECM terms are a way of capturing adjustments in a dependent variable which depend not on the levels of some explanatory variable, but on the extent to which an explanatory variable deviated from an equilibrium relationship, of the form

$$y^* = \theta^* x^*$$

with the dependent variable.

$$C_t^* = AY_t$$

consume proportional to income. C_t^* consumer target.

$A \equiv$ marginal propensity to consume (unobserved). We suppose A is observable, taking logs

$$\log C_t^* = \log A + \log Y_t$$

$$c_t^* = a + y_t$$

ECM

$$\Delta c_t = \alpha \Delta c_t^* + \gamma \left(c_{t-1}^* - c_{t-1} \right) + u_t$$

the rate of growth of aggregate consumption as a function of target consumption.

c_{t-1}^* target consumption at time $t-1$, c_{t-1} effective consumption at time $t-1$

$$\begin{aligned} \Delta c_t^* &= \Delta a + \Delta y_t \\ &= \Delta y_t \end{aligned}$$

$$\Delta c_t = \alpha \Delta y_t + \gamma \left(a + y_{t-1} - c_{t-1} \right) + u_t$$

$$\Delta c_t = a\gamma + \alpha \Delta y_t + \gamma \left(y_{t-1} - c_{t-1} \right) + u_t$$

Long-run equilibrium:

$$E\left(\Delta c_t\right) = a\gamma + \gamma E\left(y_{t-1} - c_{t-1}\right) = 0$$

$$c^* = a + y^*$$

$$C^* = AY^*$$