Chapter 3
Single Equation Dynamic Linear Models – an Introduction

Artur Silva Lopes, March 2022

ISEG–ULisboa

1 Introduction

In economics the adequate modelling of many phenomena requires that in the same equation there are variables observed in distinct time periods, lagged variables, with *lags*. For instance, consumers' behavior is not exclusively determined by their current income. Instead it results also from habits from the past, from expectations about the future, etc.

The need to include lagged variables derives mainly from:

a) habit persistence;

b) restrictions (technological, institutional, etc.);

c) expectations effects of economic agents, etc..

More generally: we have no reason to assume that data come from a state of equilibrium in the economy. It seems better to view the economy, in each period, as trying to adjust to shocks from the past, but simultaneously receiving new shocks that push her away from equilibrium.

2 Systematic Dynamics

2.1 DL(s) models and multipliers

DL(s) models are "*distributed lag*" models: the effects of changes in variables are distributed over several time periods. For instance the DL(2):

$$y_t = 0.4x_t + 1.0x_{t-1} + 0.6x_{t-2} + u_t.$$

If at moment t we give a temporary or transitory unit variation to x, with the errors silenced, what happens to y? The following path for (the expectation) of

y:
$\Delta y_t = y_t - y_{t-1} = 0.4$,
$y_{t+1} - y_{t-1} = 1$,
$y_{t+2} - y_{t-1} = 0.6$,
$y_{t+3} - y_{t-1} = 0$, $\quad \ldots$
0.4 is the impact or short run multiplier. The total or cumulated effect, the sum of the variations, called *long run multiplier*, is 2.0.

This kind of analysis is called the systematic dynamics analysis.

$s = 2$ is the lag length or the maximum lag or the lag truncation parameter.

The previous model can be written as

$$y_t = B(L)x_t + u_t, \quad \text{with} \quad B(L) = 0.4 + L + 0.6L^2,$$

L denoting the usual lag operator ($L^k y_t = y_{t-k}$).

These models can be extended, with the errors following an ARMA stationary process, $\phi(L)u_t = \theta(L)\epsilon_t$:

$$y_t = \frac{B(L)}{A(L)}x_t + \frac{\theta(L)}{\phi(L)}\epsilon_t,$$

and this is called a transfer function model. We will not study them.

2.2 ADL (ARDL) models: introduction

The most important models: ADL (*autoregressive distributed lag*) which, besides a distributed lag component, (DL), also have an autoregressive part (A or AR). The ADL(1,1):

$$y_t = \mu + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t, \quad \epsilon_t \sim iid(0, \sigma^2). \text{ Or}$$

$$\underbrace{(1 - \alpha_1 L)}_{A(L)} y_t = \mu + \underbrace{(\beta_0 + \beta_1 L)}_{B(L)} x_t + \epsilon_t,$$

$$A(L)y_t = \mu + B(L)x_t + \epsilon_t. \tag{1}$$

Suppose that $\Delta x_t = 1$ but $x_{t+1}, x_{t+2}, \ldots, x_{t+j}$ are held fixed. With $\epsilon_t \equiv 0, \forall t$, which is the trajectory of (the expectation of) y?
$\Delta y_t = \beta_0$ (short run multiplier),
$y_{t+1} - y_{t-1} = \beta_1$ (via DL) $+ \alpha_1\beta_0$ (via AR),
$y_{t+2} - y_{t-1} = \beta_0\alpha_1^2 + \beta_1\alpha_1, \ldots$ (via AR and the DL channel becames mute).

More sophisticated way: solve (1) with respect to y_t, multiplying by $A(L)^{-1}$:

$$y_t = A(L)^{-1}\mu + A(L)^{-1}B(L)x_t + A(L)^{-1}\epsilon_t,$$

deleting the ϵ_t component

$$y_t = A(1)^{-1}(\mu) + (1 - \alpha_1 L)^{-1}B(L)x_t,$$

since μ is a constant ($A(1) = 1 - \alpha_1 - \ldots - \alpha_r$).

$$\begin{aligned} y_t &= A^{-1}(1)\mu + (1 + \alpha_1 L + \alpha_1^2 L^2 + \ldots)(\beta_0 + \beta_1 L)x_t, \\ &= A^{-1}(1)\mu + [\beta_0 + (\beta_1 + \alpha_1\beta_0)L + (\beta_1\alpha_1 + \alpha_1^2\beta_0)L^2 + \ldots]x_t, \\ &= A^{-1}(1)\mu + \beta_0 x_t + (\beta_1 + \alpha_1\beta_0)x_{t-1} + (\beta_1\alpha_1 + \alpha_1^2\beta_0)x_{t-2} + \ldots, \end{aligned}$$

Dynamic multipliers: $\partial y_{t+j}/\partial x_t$. Viewed as a function of the lag, j, these multipliers are the impulse response function because they provide the response of y to a single impulse, transitory and unitary, of x. Sometimes they are graphically represented.

Long-run multiplier. Or total, or equilibrium, which is the sum of the series

$$\text{LRM} = \sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial x_t},$$

it measures the total accumulated effect in y resulting from the unit transitory variation of x. In the previous DL(2) model: $B(L) = 0.4 + 1.L + 0.6L^2 \Rightarrow B(1) = 0.4 + 1 + 0.6 = 2$.

In the case of the ADL

$$\sum_{j=0}^{\infty} \frac{\partial y_{t+j}}{\partial x_t} = \beta_0 + (\beta_0\alpha_1 + \beta_1) + (\beta_0\alpha_1^2 + \beta_1\alpha_1) + \ldots .$$

Provided $|\alpha_1| < 1$ the series is convergent and the total multiplier is

$$\lambda = \frac{\beta_0 + \beta_1}{1 - \alpha_1},$$

That is,

$$\text{LRM} = \lambda = \frac{B(1)}{A(1)}.$$

Alternatively: suppose that the economy is in an (hypothetical) state of (static) stationary equilibrium: all the variables are in their state of equilibrium, the mean in the stationary case:

$$y^* = \mathsf{E}(y_t), \quad x^* = \mathsf{E}(x_t), \quad \text{and} \quad \epsilon^* = \mathsf{E}(\epsilon_t) = 0.$$

Then

$$y^* = \mu + \alpha_1 y^* + \beta_0 x^* + \beta_1 x^*,$$

that is, the stationary (static) equilibrium solution is

$$y^* = \frac{\mu}{1-\alpha_1} + \frac{\beta_0 + \beta_1}{1-\alpha_1} x^*.$$

If x now has a unit *permanent* variation, in the new equilibrium and provided the system is stable we will have

$$\Delta y^* = \frac{\beta_0 + \beta_1}{1-\alpha_1} = \lambda,$$

that is, $\Delta y^* = LRM$. The LRM represents the equilibrium variation of y, its long-run variation, corresponding to an unit permanent variation of the equilibrium value of x. That is, the LRM is

$$\text{LRM} = \lim_{j\to\infty} \left(\frac{\partial y_{t+j}}{\partial x_t} + \frac{\partial y_{t+j}}{\partial x_{t+1}} + \ldots + \frac{\partial y_{t+j}}{\partial x_{t+j}} \right).$$

Stability conditions. The condition $|\alpha_1| < 1$ is a stability condition of the difference equation. When $|\alpha_1| < 1$ the effects of the transitory variation of x eventually become negligible, they go to zero, and the economy returns to a new equilibrium, with a finite value for y (even when the variation in x is permanent).

If $|\alpha_1| > 1$ the system would be explosive. If $\beta_0 > 0, \beta_1 > 0$ and $\alpha_1 > 1$, when $t \to \infty$, $y \to \infty$.

$\alpha_1 = 1$ (and $\alpha_1 = -1$) is not reasonable too: after any number of periods, y will continue varying $\beta_0 + \beta_1$ units, it does not converge (and with $\alpha_1 = -1$ the behaviour is oscillatory).

The stability analysis has to do only with the autoregressive component of the model.

Let us now generalize to a general ADL, with orders r and s (ADL(r, s)), but with a single exogenous variable

$$y_t = \mu + \alpha_1 y_{t-1} + \ldots + \alpha_r y_{t-r} + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_s x_{t-s} + \epsilon_t,$$

that is,

$$A(L)y_t = \mu + B(L)x_t + \epsilon_t,$$

with

$$A(L) = 1 - \alpha_1 L - \ldots - \alpha_r L^r \quad \text{and} \quad B(L) = \beta_0 + \beta_1 L + \ldots + \beta_s L^s,$$

the autoregressive and the distributed lag polynomials, respectively.

The necessary but not sufficient condition for stability is that

$$\sum_{i=1}^{r} \alpha_i < 1 \;\;\Leftrightarrow\;\; \sum_{i=1}^{r} \alpha_i - 1 < 0 \;\;\Leftrightarrow\;\; -A(1) < 0,$$

which, in the ADL(1,1) model is $\alpha_1 < 1$. Actually, this the only condition that we will care about because it is the most important in economics.

Necessary and sufficient condition for stability: the values of z: $A(z) = 0 \Leftrightarrow 1 - \alpha_1 z - \alpha_2 z^2 - \ldots - \alpha_r z^r = 0$, represented with z_i, $i = 1, 2, \ldots, r$, i.e., the roots of $A(z)$, must be outside the unit circle, $|z_i| > 1, \forall i$.

In practice, we will care only with the necessary condition. Provided there is stability, in this general case it emerges that

$$\mathrm{LRM} = \lambda = \frac{B(1)}{A(1)} = \frac{1 + \beta_0 + \ldots + \beta_s}{1 - \alpha_1 - \ldots - \alpha_r}.$$

Notes: a) if all variables are logarithmized, instead of multipliers we have elasticities (short- and long-run); b) in the case of the log-lin models we will have the semi-elasticities.

Other concepts: lag weights, average and median lags, etc. .

3 OLS estimation and linear transformations

Important problem: precise estimation of the DL and ADL coefficients.

3.1 Colinearity problems

Typically, macroeconomic time series are highly autocorrelated (positively): the correlation coefficients between x_t and its lagged values (and between them) are high (the same with y_t). Colinearity problems then follow: lag structures are estimated poorly, i.e., imprecisely. Moreover: the ses become inflated $\Rightarrow$ t-ratios become deflated $\Rightarrow$ individual tests have low power.

Moreover, uncertainty about the coefficients also increases: when a lag is excluded the estimates of the remaining coefficients usually change a lot.

3.2 Linear transformations

To reduce the size of the problem: the model will be reparameterized (without any restriction) using linear transformations.

Proposition: the OLS estimator of linear regression models is invariant to linear transformations. That is: it does not matter whether we estimate the initial model with OLS or the reparameterized model and get the original estimates using the linear transformations of the reparameterization. But the estimation of this last model can be advantageous in that it may reduce the colinearity problems of the original model.

Proof: consult the text.

The advantage is that in the reparameterized model there could be (almost) orthogonality between the regressors, that is, the colinearity of the original model can be (almost) eliminated by means of the linear transformations.

3.3 A simple example: application to the DL(1) model

$$y_t = \delta_0 x_t + \delta_1 x_{t-1} + u_t \tag{2}$$

Problem: x_t highly autocorrelated, implying:

- the estimators' ses are very high;
- therefore, (at least) one of the coefficients may appear non statistically significant $\Rightarrow$ incorrect exclusion from the model;
- this provokes a large variation in the estimate of the remaining coefficient.

Solution: replace x_t with $x_{t-1} + \Delta x_t$ (or summing and subtracting $\delta_0 x_{t-1}$ to the right hand side of the equation),

$$y_t = \delta_0 \Delta x_t + (\delta_0 + \delta_1) x_{t-1} + u_t, \tag{3}$$

a) Δx_t and x_{t-1} will be weakly correlated (they will be almost orthogonal), and the estimate of $\delta_0 + \delta_1$ is insensitive to the presence (or absence) of Δx_t in (from) the regression;

b) we obtain immediately an estimate of the LRM ($\delta_0 + \delta_1$) through the coefficient of x_{t-1}, as well as its $se \Rightarrow$ we can perform immediately a significance test of that multiplier.

We know that the OLS estimates are exactly equal to those of (2). But we must use preferentially (3) due to its advantage.

Empirical example. The consumption function for the portuguese economy. Model:

$$LPC_t = \alpha + \delta_0 LDI_t + \delta_1 LDI_{t-1} + u_t,$$

where, LPC and LDI are the logarithms of private (household) consumption and disposable income, respectively. With annual data for 1960 to 1995:

$$\begin{array}{llll} \widehat{LPC}_t = & 0.526+ & 0.556 LDI_t + & 0.323 LDI_{t-1} \\ & (0.163) & (0.296) & (0.283) \end{array}$$

and at the 5% level none of the individual coefficients is significant, the t-ratios being 1.878 and 1.140. $\widehat{LRE} = 0.879 (= 0.556 + 0.323)$ and the significance test demands some work (to obtain $se(\widehat{LRE})$).

Excluding LDI_{t-1} (since it is the less significant):

$$\begin{array}{lll} \widehat{LPC}_t = & 0.416+ & 0.895 LDI_t, \\ & (0.130) & (0.021) \end{array}$$

The estimated short-run elasticity changed a lot, provoking uncertainty. But the estimate of the LRE (which is the same) changed little.

The linear reparameterization produces

$$\begin{array}{llll} \widehat{LCP}_t = & 0.526+ & 0.556\Delta LRD_t+ & 0.879 LRD_{t-1}, \\ & (0.163) & (0.296) & (0.025) \end{array}$$

and the problem becomes much less serious: only the coefficient of ΔLRD_t does not appear statistically significant at 5%, with a t-ratio equal to 1.878.The estimate of the LRE is, immediately, 0.879, and it is highly significant ($t = 34.71$). And now even ΔLRD_t appears almost statistically significant at the 5% level.

3.4 Generalization: application to DL(s) models

Consider the DL(s) model, without the intercept, to simplify

$$\begin{aligned} y_t = & \ \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_s x_{t-s} + u_t \\ = & \ B(L) x_t + u_t, \end{aligned}$$

with $B(L) = \beta_0 + \beta_1 L + \ldots + \beta_s L^s$ an order s polynomial in L. The linear transformation requires that we first consider a proposition about polynomial decompositions.

Proposition. Consider a polynomial of order p in L, say

$$a(L) = \sum_{j=0}^{p} a_j L^j = a_0 + a_1 L + a_2 L^2 + ... + a_p L^p,$$

and define the coefficients $c_0 = a_0$, $c_i = -\sum_{j=i+1}^{p} a_j$, $i = 1, ..., p-1$ and $c_p = 0$. Then, one can write:

$$a(L) = a(1)L + \sum_{i=0}^{p-1} c_i L^i (1 - L), \tag{4}$$

that is,

$$a(L) = a(1)L + c(L)(1 - L), \tag{5}$$

where $c(L)$ is a polynomial of order $p - 1$ in L.

Applying this result to the DL(s) model we get ($\Delta = 1 - L$):

$$\begin{aligned} y_t = & \; B(L)x_t + u_t \\ = & \; [B(1)L + \sum_{i=0}^{s-1} \delta_i L^i (1 - L)]x_t + u_t \\ = & \; B(1)x_{t-1} + \sum_{i=0}^{s-1} \delta_i \Delta x_{t-i} + u_t, \end{aligned}$$

where $\delta_0 = \beta_0$ (i.e., the SRM is not affected), $\delta_i = -\sum_{j=i+1}^{s} \beta_i, \; i = 1, 2, ..., s-1$ and $\Delta = 1 - L$. Estimating this model is equivalent to estimating the original model. However, this last model provides several benefits:

a) it allows us obtaining immediately an estimate of the long-run multiplier (the coefficient of x_{t-1}) and of its estimator'standard error (se);

b) hence, we can test immediately the significance of that parameter

c) the colinearity problems of the initial model should become much reduced, because both the correlation between x_t and the lags of Δx_t as well as the correlations between these must be much smaller than those between the regressors of that model.

Note: in the reparameterized model (besides the SRM) the original coefficients become mixed. To retrieve them we need to solve the equations that connect them to the coefficients of the reparameterized model.

In models with more variables these reparameterizations are even more necessary (and more useful). Only the lagged level of each variable will remain in the model; all other regressors will be first differences and therefore they will be much less correlated between them and with those lagged variables than the original regressors.

4 The ADL(1,1) model and the ECM

This type of model will be very important later, in the multivariate empirical analysis of non-stationary time series (but becoming stationary when they are differenced once, i.e., integrated of order 1, I(1)).

4.1 The generality of the ADL(1,1) model

Purpose: to show that ADL models are very general, considering the example of the ADL(1,1), with only one variable assumed as exogenous:

$$y_t = \mu + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t, \ \epsilon_t \sim iid(0, \sigma^2), \quad \text{with} \ |\alpha_1| < 1 \quad (6)$$

The following models are particular cases of the ADL(1,1), obtained imposing restrictions on their coefficients:

1) simple, static regression $y_t = \mu + \beta_0 x_t + \epsilon_t$, with $\alpha_1 = \beta_1 = 0$;

2) AR(1) model, with $\beta_0 = \beta_1 = 0$;

3) model in first differences, $\Delta y_t = \mu + \beta_0 \Delta x_t + \epsilon_t$, with $\beta_1 = -\beta_0$ and $\alpha_1 = 1$ (non stable);

4) leading indicator: $y_t = \mu + \beta_1 x_{t-1} + \epsilon_t$, with $\alpha_1 = \beta_0 = 0$;

5) the partial adjustment model, $y_t = \mu + \alpha_1 y_{t-1} + \beta_0 x_t + \epsilon_t$, with $\beta_1 = 0$;

6) static model with AR(1) errors, the common factor model, $y_t = \theta + \beta_0 x_t + u_t$, $u_t = \alpha_1 u_{t-1} + \epsilon_t$, with $\mu = \theta(1 - \alpha_1)$ and $\beta_1 = -\alpha_1 \beta_0$;

7) DL(1) model, $y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t$, with $\alpha_1 = 0$;

8) "*dead start*" model ($\approx$ adaptative expectations model), $y_t = \mu + \alpha_1 y_{t-1} + \beta_1 x_{t-1} + \epsilon_t$, with $\beta_0 = 0$.

Hendry (1995) investigates the properties of all these models when the DGP or "true model" is the ADL(1,1): estimators'inconsistency and several misspecification symptons emerge ("residual autocorrelation", parameter instability, etc.).

The most important parameterization of the ADL(1,1) is the one of error correction:

$$\Delta y_t = (\alpha_1 - 1)(y_{t-1} - \lambda_0 - \lambda_1 x_{t-1}) + \beta_0 \Delta x_t + \epsilon_t,$$

where $\lambda_0 = \mu/(1 - \alpha_1)$ and $\lambda_1 = (\beta_0 + \beta_1)/(1 - \alpha_1)$, but where there isn't any restriction.

If the restriction $\alpha_1 + \beta_0 + \beta_1 = 1$ is imposed, the "homogeneous ECM" results, given by

9) $\Delta y_t = (\alpha_1 - 1)(y_{t-1} - \lambda_0 - x_{t-1}) + \beta_0 \Delta x_t + \epsilon_t$, where the long-run multiplier is one.

4.2 Linear transformations of the ADL(1,1) model

Consider an ideal, long-run situation, of a stationary equilibrium, all the variables at their equilibrium values, their means.

Recall that $y^* = \underbrace{\frac{\mu}{1-\alpha_1}}_{\lambda_0} + \underbrace{\frac{\beta_0 + \beta_1}{1-\alpha_1}}_{\lambda_1} x^*$.

Now, subtracting y_{t-1} to both members of the ADL(1,1) from (6) one gets

$$\Delta y_t = \mu + (\alpha_1 - 1) y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t, \tag{7}$$

Summing and subtracting $\beta_0 x_{t-1}$ to the right hand side, one gets the *"Bardsen form"*:

$$\Delta y_t = \mu + \underbrace{(\alpha_1 - 1)}_{-A(1)} y_{t-1} + \beta_0 \Delta x_t + \underbrace{(\beta_0 + \beta_1)}_{B(1)} x_{t-1} + \epsilon_t, \tag{8}$$

and to estimate the LRM one simply takes the symmetrical of the quotient between the estimate of the coefficient of x_{t-1} and that of y_{t-1}.

If, on the other hand, we sum and subtract $\beta_1 x_t$ to the right hand side of(7), we get

$$\Delta y_t = \mu + (\alpha_1 - 1) y_{t-1} - \beta_1 \Delta x_t + (\beta_0 + \beta_1) x_t + \epsilon_t, \tag{9}$$

where, to get the estimate of the LRM, one simply divides the symmetrical of the estimate of the coefficient of x_t by the coefficient of y_{t-1}.

The "ECM form" is obtained from (8):

$$\Delta y_t = (\alpha_1 - 1) \left[y_{t-1} - \frac{\mu}{1 - \alpha_1} - \frac{\beta_0 + \beta_1}{1 - \alpha_1} x_{t-1} \right] + \beta_0 \Delta x_t + \epsilon_t, \tag{10}$$

that is,

$$\Delta y_t = (\alpha_1 - 1)[y_{t-1} - \lambda_0 - \lambda_1 x_{t-1}] + \beta_0 \Delta x_t + \epsilon_t. \tag{11}$$

In "Bewley's form":

$$y_t = \gamma\mu - \gamma\alpha_1 \Delta y_t + \gamma(\beta_0 + \beta_1) x_t - \gamma\beta_1 \Delta x_t + \gamma\epsilon_t, \tag{12}$$

the LRM is the coefficient of x_t. But OLS is inconsistent in this equation. Why?

4.3 The error correction model

Let's get back to (11) or (10) and notice that the term inside brackets measures the error or deviation from equilibrium or desequilibrium from the previous period. These are short-run dynamic equations but $(\Delta)y$ is driven also by the long-run equilibrium relation. Even if there are no shocks for a long time period, with

$\epsilon_t \equiv 0$ and $\Delta x_t \equiv 0$ for many t, Δy will not become zero unless the equilibrium is attained, that is, until $y = \lambda_0 + \lambda_1 x$.

The coefficient $\alpha_1 - 1$ measures the short-run adjustment speed (to the equilibrium), $(\alpha_1 - 1)[y_{t-1} - \lambda_0 - \lambda_1 x_{t-1}]$ is the error correction term and the expression inside brackets is the error correction mechanism.

When the stability condition is satisfied, $|\alpha_1| < 1$, the inequality $-2 < \alpha_1 - 1 < 0$ holds, that is, the adjustment coefficient is negative, but it is larger than -2.

Suppose that in $t-1$ the equilibrium error was positive, i. e., that $y_{t-1} > \lambda_0 + \lambda_1 x_{t-1}$. Then, since $\alpha_1 - 1 < 0$, *ceteris paribus*, in the following period (t) y will show a tendency to return to its equilibrium value because the functioning of the mechanism implies that $\Delta y_t < 0$; that is, if there are no shocks $\Delta y_t < 0$, so that there is a return to equilibrium.

Conversely, if in $t-1$, $y_{t-1} < \lambda_0 + \lambda_1 x_{t-1} \Rightarrow \Delta y_t > 0$.

Hence the model contains a negative *feedback* effect trying to correct disequilibria from previous periods to reestablish the long-run equilibrium relationship.

To test for the presence of an error correction mechanism:

$$H_0 : \alpha_1 - 1 = 0 \Leftrightarrow -A(1) = 0, \quad vs. \quad H_1 : \alpha_1 - 1 < 0 \Leftrightarrow -A(1) < 0,$$

the rejection of $H_0 \Leftrightarrow$ evidence on the presence of an ECM.

For estimation, the preferred form is the Bardsen one:

a) since it provides promptly an estimate for the adjustment coefficient $(\alpha_1 - 1)$;

b) because the t-ratio of that coefficient allows testing immediately the presence of an ECM;

c) because it allows obtaining easily an estimate for the LRM and

d) because it is preferable to the initial ADL(1,1) since colinearity is reduced.

5 The ECM with non-stationary variables

Many macroeconomic time series are non-stationary. But the ECM is still an interesting model:

a) the long-run equilibrium relation is now between non-stationary variables;

b) the ECM incorporates also a "*steady growth*" solution;

c) the ECM allows the inclusion of growth effects: the level of a variable depending on the level of another but also on the growth rate of this last one.

For further details on this subject you must read the text.

6 The ADL(r, s) model and the ECM

Purpose: to generalize the reparameterization as ECMs of more general ADL models, ADL(r, s).

6.1 The reparameterization as ECM of ADL(r, s) models

The proposition about polynomial decompositions can be applied to autoregressive polynomials too.

Consider an ADL(r, s) with observations of a single variable, assumed as exogenous and without intercept to simplify:

$$y_t = \sum_{i=1}^{r} \alpha_i y_{t-i} + \sum_{i=0}^{s} \beta_i x_{t-i} + \epsilon_t, \quad \epsilon_t \sim \text{iid}(0, \sigma^2), \tag{13}$$

that is,

$$A(L)y_t = B(L)x_t + \epsilon_t, \tag{14}$$

with $A(L) = 1 - \sum_{i=1}^{r} \alpha_i L^i$ and $B(L) = \sum_{i=0}^{s} \beta_i L^i$.

Using the decomposition of the autoregressive polynomial, we can write

$$A(L) = A(1)L + (1-L)D(L)$$

where $D(L)$ has an order which is one unit less than the initial

$$D(L) = 1 - \sum_{i=1}^{r-1} \delta_i L^i = 1 - \sum_{i=1}^{r-1} (-\sum_{j=i+1}^{r} \alpha_j) L^i.$$

On the other hand, using the decomposition of the $B(L)$ polynomial, we get

$$B(L) = B(1)L + (1-L)G(L),$$

where

$$G(L) = \sum_{i=0}^{s-1} \gamma_i L^i = \gamma_0 + \sum_{i=0}^{s-1} (- \sum_{j=i+1}^{s} \beta_j) L^i.$$

Using these two polynomial decompositions in (14) one gets

$$[A(1)L + (1 - L)D(L)]y_t = [B(1)L + (1 - L)G(L)]x_t + \epsilon_t,$$

that is (since $(1 - L) = \Delta$),

$$A(1)y_{t-1} + (1 - \sum_{i=1}^{r-1} \delta_i L^i)\Delta y_t = B(1)x_{t-1} + \sum_{i=0}^{s-1} \gamma_i \Delta x_{t-i} + \epsilon_t,$$

from which one easily gets the Bardsen form:

$$\Delta y_t = -A(1)y_{t-1} + \sum_{i=1}^{r-1} \delta_i \Delta y_{t-i} + B(1)x_{t-1} + \sum_{i=0}^{s-1} \gamma_i \Delta x_{t-i} + \epsilon_t, \quad (15)$$

an ADL$(r-1, s-1)$ in Δy_t and Δx_t augmented with y_{t-1} and x_{t-1}, the only two regressors in level form.

The ECM representation is derived isolating the $-A(1)$ term:

$$\Delta y_t = -A(1)\left[y_{t-1} - \frac{B(1)}{A(1)} x_{t-1}\right] + \sum_{i=1}^{r-1} \delta_i \Delta y_{t-i} + \sum_{i=0}^{s-1} \gamma_i \Delta x_{t-i} + \epsilon_t, \quad (16)$$

that is, with ϕ denoting the ajustment coefficient, $\phi = -A(1)$ and $\lambda = B(1)/A(1) =$LRM,

$$\Delta y_t = \phi\left[y_{t-1} - \lambda x_{t-1}\right] + \sum_{i=1}^{r-1} \delta_i \Delta y_{t-i} + \sum_{i=0}^{s-1} \gamma_i \Delta x_{t-i} + \epsilon_t, \qquad (17)$$

where $\phi < 0$, otherwise there is no ECM, i.e., provided the necessary (but not sufficient) condition for stability is satisfied (that is, provided $\sum_{i=1}^{r} \alpha_i < 1$, that is, $A(1) > 0$).

Again, the form usually employed for estimation (with OLS) is the Bardsen one.

Notice again the advantages of (15):

1) it allows obtaining immediately an estimate of the adjustment coefficient, $-A(1) = \phi$;

2) the t-ratio for this coefficient may be used also immediately to perform the t-ECM test (*t-ECM*), that is, to test for the presence of the error correction mechanism ($H_0 : \phi = 0$ $vs.$ $H_1 : \phi < 0$);

3) it allows obtaining easily an estimate for the LRM and

4) it allows reducing the problems arising from colinearity of the initial model.

Extending this to the case of several (k) exogenous variables requires only small adaptations:

- $\mathbf{x}_t$ now represents a vector, $k \times 1$, $\mathbf{x}'_t = [x_{t1} \;\; x_{t2} \;\; \ldots \;\; x_{tk}]$, with lag lenths given by $s_1, s_2, ..., s_k$. The model is still an ADL(r, s), with $s = \max\{s_1, s_2, ..., s_k\}$.

- the $\boldsymbol{\beta}_i$'s are now vectors, $k \times 1$, of lag coefficients, that is,

$$\boldsymbol{\beta}'_i = [\beta_{i1}\ \beta_{i2}\ ...\ \beta_{ik}],\ i = 0, 1, ...s.$$

- (14) is now written as

$$A(L)y_t = \mathbf{B}(L)'\mathbf{x}_t + \epsilon_t, \tag{18}$$

where $\mathbf{B}(L) = \sum_{i=0}^{s} \boldsymbol{\beta}_i L^i$ represents a vector, $k \times 1$, of polynomials in L, that is, $\mathbf{B}(L)' = [B_1(L)\ B_2(L)\ \ldots\ B_k(L)]$, with $B_j(L), j = 1, \ldots, k$.

- With $\boldsymbol{\lambda}$ denoting the vector, $k \times 1$, of long-run multipliers, $\boldsymbol{\lambda} = A(1)^{-1}\mathbf{B}(1) = -\phi^{-1}\mathbf{B}(1)$, the ECM representation is:

$$\Delta y_t = \phi[y_{t-1} - \boldsymbol{\lambda}'\mathbf{x}_{t-1}] + \sum_{i=1}^{r-1} \delta_i \Delta y_{t-i} + \sum_{i=0}^{s-1} \boldsymbol{\gamma}'_i \Delta \mathbf{x}_{t-i} + \epsilon_t. \tag{19}$$

- the (long-run) stationary equilibrium is

$$y^* = \boldsymbol{\lambda}'\mathbf{x}^* = \lambda_1 x_1^* + \lambda_2 x_2^* + \ ... \ + \lambda_k x_k^*.$$

Since there are now several exogenous variables, the benefits of the Bardsen reparameterization in terms of reduction of the colinearity problems are even stronger.

6.2 The general to specific (GTS) modelling

GTS (or GETS) is a modelling strategy to get an adequate dynamic model that consists in starting work with a rather general model — an ADL(r, s) with high orders —, and proceed "testing down", until one gets a simpler but satisfactory model.

This strategy encompasses the following stages or steps:

1. Start with a dynamic model with a high order, an ADL(r, s) preferably in the Bardsen form, that is both coherent with the equilibrium relation given by economic theory and that does not impose any restrictions on the short run dynamics.

2. Simplify the model excluding non significant regressors or imposing other restrictions that do not violate data and that do not imply that symptoms of specification errors appear.

3. Final assessment of the model based in economic theory and in misspecification tests.

6.3 The specific to general approach

For details read the text.

7 Empirical example

The data are again for the aggregate consumption function for the portuguese economy, from 1965 to 1995. Recall the conventions:
$\text{DLC}_t = \Delta LC_t = LCP_t - LCP_{t-1}$, first difference of the log of household consumption,
$\text{DLR}_t = \Delta LRD_t = LRD_t - LRD_{t-1}$, first difference of the log of disposable income,

DINF$_t = INF_t - INF_{t-1}$ first difference of inflation and
DLS$_t = LSR_t - LSR_{t-1}$ the first difference of an index of real wages.

Since the data are annual and the sample is short: started with an ADL(3,3) in the levels of the variables, parameterized in the Bardsen form. Some additional ideas about the path:

a) Since there is more than a single possible path, one can arrive to distinct final specifications.

b) The preferred specification tests were the serial correlation tests. Despite being non significant, the constant term was maintained, not only because its presence is common, but also because its removal would imply a significant deterioration in the quality of forecasts.

c) Despite non significant at the 5% level, INF_{t-1} was retained because its deletion would imply the exclusion from the long run equilibrium relationship.

Concerning the significance of the regressors Equation 1 is bleak, with only one significant at 5%. Note, however, the global significance almost at the 1% level, and mainly the absence of any visible symptoms of specification problems.

```
                                  Equation   1
                                  =============
                 Method of estimation = Ordinary Least Squares
Dependent variable: DLC
                 R-squared = .831171
              LM het. test = 2.49818 [.114]
Breusch/Godfrey LM: AR/MA1 = .074112 [.785]
Breusch/Godfrey LM: AR/MA3 = 1.94985 [.583]
            Ramsey's RESET2 = 1.42562 [.258]
          F (zero slopes) = 3.93853 * [.011]
```

Variable	Estimated Coefficient	Standard Error	t-statistic		P-value
C	.065127	.128018	.508735		[.620]
LCP(-1)	-.287277	.254321	-1.12958		[.281]
LRD(-1)	.261714	.240360	1.08884		[.298]
LSR(-1)	.083553	.087219	.957966		[.357]
INF(-1)	-.818099E-03	.280483E-02	-.291675		[.776]
DLC(-1)	.395331	.356604	1.10860		[.289]
DLC(-2)	.154389	.399052	.386890		[.706]
DLR	.380204	.167385	2.27144	*	[.042]
DLR(-1)	-.083260	.218713	-.380681		[.710]
DLR(-2)	-.290114	.202744	-1.43094		[.178]
DLS	.286027	.247846	1.15405		[.271]
DLS(-1)	.027639	.299438	.092304		[.928]
DLS(-2)	-.767700E-02	.263578	-.029126		[.977]
DINF	.875192E-03	.270116E-02	.324006		[.752]
DINF(-1)	.729020E-04	.234304E-02	.031114		[.976]
DINF(-2)	.353931E-03	.141999E-02	.249248		[.807]

Being careful to keep the same sample*, a first simplification F-test, of joint significance of all the 2-period lagged differenced regressors was performed: $F = 3.46$, with $p - value \approx 0.042 \Rightarrow$ rejection of H_0.

Note that the rejection must be due to ΔLR_{t-2}, which in not far from significant. Deleting from the previous test the nullity of that coefficient, the F-statistic changes to 0.76, with a $p - value \approx 0.538 \Rightarrow$ non rejection of the joint nullity of the 3 coefficients $\Rightarrow$ the variables ΔLC_{t-2}, ΔLS_{t-2} and ΔINF_{t-2} are removed from the model.

. . .

*This caution is important since the exclusion of the second lag of all variables liberates one observation and estimating the model with the restrictions can begin in 1967. This should not be done so that the sample that provides the SSR entering the F statistic is maintained.

Equation 6

=============

Dependent variable: DLC

R-squared = .814592

LM het. test = .140255 [.708]

Breusch/Godfrey LM: AR/MA1 = .178087E-02 [.966]

Breusch/Godfrey LM: AR/MA3 = 3.53389 [.316]

Chow test = 2.27653 [.093]

Ramsey's RESET2 = .661594 [.426]

C	.048776	.082447	.591607		[.560]
LCP(-1)	-.266245	.071495	-3.72395	**	[.001]
LRD(-1)	.243360	.065785	3.69935	**	[.001]
LSR(-1)	.095209	.039934	2.38416	*	[.027]
INF(-1)	-.919432E-03	.585899E-03	-1.56927		[.132]
DLC(-1)	.456655	.135524	3.36956	**	[.003]
DLR	.338044	.105065	3.21748	**	[.004]
DLR(-2)	-.247299	.110304	-2.24199	*	[.036]
DLS	.225807	.079575	2.83767	**	[.010]

The equation does not seem to present any specification problems.

In terms of goodness of fit, and considering the number of parameters, the final model is better than the initial.

Goodness of fit statistics (1968-95)

	initial model	final model
R^2	0.831	0.814
$\bar{R}^2$	0.620	0.735
AIC	−67.99	−73.60
BIC	−57.33	−67.60
$\widehat{\sigma}^2$	0.00034	0.00024

Is there evidence on the presence of an ECM? With $\phi = -A(1)$ the coefficient of LCP_{t-1}, $t_\phi = t_{MCE} = -3.72$ which is very significant by conventional criteria. The estimated coefficient is negative but small, in absolute value (−0.266): to return to equilibrium it is estimated that it will take approximately 4 years, which represents a slow adjustment.

The estimates for the long-run elasticities: $\lambda^{LR} = -\frac{0.24336}{-0.266244} = 0.914$, and $\lambda^{LS} = -\frac{0.095209}{-0.266245} = 0.358$. Estimate of the long-run semi-elasticity with respect to inflation: $\lambda^{INF} = -\frac{0.00092}{-0.266245} = -0.00034$. Intercept estimate in the long-run equilibrium relation: $\lambda_0 = -\frac{0.048774}{-0.266245} = 0.183$.

Long-run equilibrium estimated relationship:

$$\widehat{LCP} = 0.183 + 0.914LRD + 0.358LSR - 0.00034INF$$

where the estimated long-run elasticity relatively to income is much larger than the one for short-run (0.338).

Finally, the estimated ECM

$$\begin{aligned}\widehat{\Delta LC}_t = \; & -0.266(LCP_{t-1} - 0.183 - 0.914LRD_{t-1} - 0.358LSR_{t-1} + 0.00034INF_{t-1}) \\ & +0.338\Delta LR_t + 0.457\Delta LC_{t-1} + 0.226\Delta LS_t - 0.247\Delta LR_{t-2},\end{aligned}$$

which has produced good forecasts for the first post sample years.