Dynamic Regression Models

Gianluca Cubadda

Università di Roma "Tor Vergata"

Interdependence

We wish to model the dynamic relationships between elements of a bivariate stochastic process $v_t \equiv (y_t, x_t)'$

Definition: Let the process v_t admit the following Vector Auto-Regressive representation of order p (VAR(p) henceforth)

$$A(L)v_t = m + \varepsilon_t$$

where $A(L) = I_2 - \sum_{j=1}^{p} A_j L^j$, m is a $2-$vector of constants, and $\varepsilon_t \equiv (\varepsilon_{1t}, \varepsilon_{2t})'$ are i.i.d. $N_2(0, \Sigma)$ innovations w.r.t. $\Omega_{t-1} = \{v_{t-i}, i = 1, 2, ...,\}$.

The conditional moments are

$$\mathsf{E}(v_t|\Omega_{t-1}) = \mathsf{E}(v_t|V_{t-1}) = m + \sum_{j=1}^{p} A_j v_{t-j},$$
$$\mathsf{Var}(v_t|\Omega_{t-1}) = \mathsf{Var}(v_t|V_{t-1}) = \Sigma,$$

where $V_{t-1} = \left\{v_{t-j}, j = 1, 2, ..., p\right\}$.

When the roots of $\det(A(z))$ are outside the unite circle, the process v_t is weakly stationary. The unconditional expected value and the autocorrelation function are then obtained as follows

$$\mathsf{E}(v_t) = \mu = A(1)^{-1} m$$
$$\mathsf{Cov}(v_t, v_{t-k}) = \Gamma(k) = \sum_{j=1}^{p} A_j \Gamma(k-j) + \mathsf{E}(\varepsilon_t v_{t-k}'),$$

where $\mathsf{E}(\varepsilon_t v_{t-k}') = 0$ if $k > 0$ and Σ if $k = 0$.

Alternative representations of a stationary VAR(p):

- Infinite-order Vector Moving Average (VMA)

$$v_t - \mu = A(L)^{-1}\varepsilon_t \equiv C(L)\varepsilon_t,$$

where $C(L) = I_2 + \sum_{j=1}^{\infty} C_j L^j$.

- The final equations

$$\text{det}(A(L))(v_t - \mu) = \text{adj}(A(L))\varepsilon_t,$$

from which it follows that both y_t and x_t admits an univariate ARMA representation of order, at most, $(2p, p)$.

The joint density of the bivariate VAR(p): The conditional joint density of $(y_t, x_t)'$ is given by

$$\mathsf{P}(y_t, x_t|\Omega_{t-1}; \theta) = \mathsf{P}(y_t, x_t|Y_{t-1}, X_{t-1}; \theta),$$

where $Y_{t-1} = \left\{y_{t-j}, j = 1, 2, ...\right\}$, $X_{t-1} = \left\{x_{t-j}, j = 1, 2, ...\right\}$, and $\theta \in \Theta \subseteq \mathbb{R}^{(4p+5)}$.

The conditional joint density can always be factorized as

$$\begin{aligned} & \mathsf{P}(y_t, x_t|Y_{t-1}, X_{t-1}; \theta) \\ = \; & \mathsf{P}(y_t|Y_{t-1}, X_t; \theta_1)\mathsf{P}(x_t|Y_{t-1}, X_{t-1}; \theta_2), \end{aligned}$$

where $(\theta_1', \theta_2')'$ is a one-to-one function of θ.

Weak Exogeneity

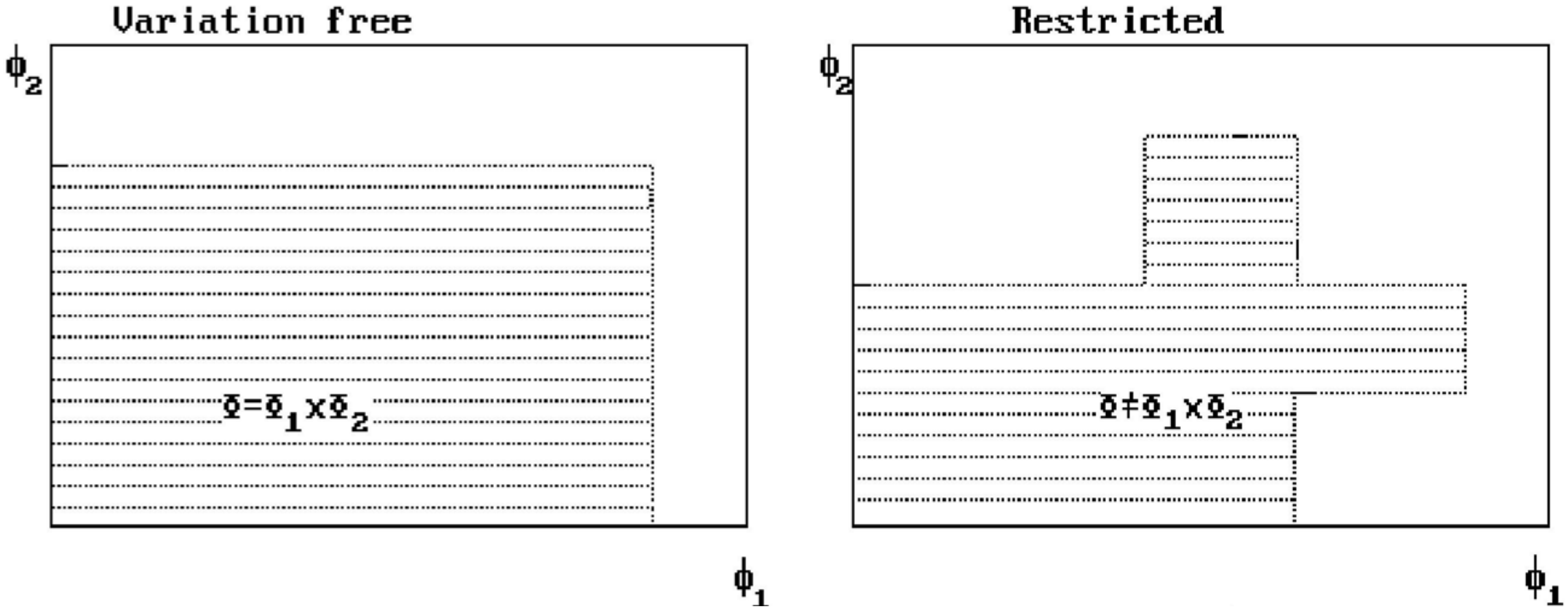
Intuition: No information loss on the parameters of interest when using a single-equation model with respect to a VAR(p) model.

Definition: The variable x is weakly exogenous for the parameters of interest, say λ, if the following conditions hold:

1. $\lambda = f(\theta_1)$ alone;

2. θ_1 and θ_2 are variation free, i.e. there are no-cross restrictions linking θ_1 and θ_2.

Parameter space restrictions



If weak exogeneity for λ holds, ML inference on λ is obtained by solving

$$\widehat{\theta}_1 = \arg\max \left\{ \prod_{t=1}^{T} \mathsf{P}(y_t, |Y_{t-1}, X_t; \theta_1) \right\},$$

and taking $\widehat{\lambda} = f(\widehat{\theta}_1) \Rightarrow$ a single-equation framework is enough.

Definition: Under weak exogeneity for parameters θ_1, we obtain the following Autoregressive Distributed Lag model of orders (m,n) (ADL(m,n) henceforth)

$$\phi(L)y_t = \delta + \beta(L)x_t + \epsilon_t,$$

where $\phi(L) = 1 - \sum_{j=1}^{m} \phi_j L^j$, $\beta(L) = \sum_{i=0}^{n} \beta_i L^i$, $\sup\{m, n\} \leq p$, ϵ_t are i.i.d. $N(0, \sigma_\epsilon^2)$ innovations w.r.t. (Ω_{t-1}, x_t), and δ is a constant.
Notice that $\theta_1 = (\delta, \phi_1, ...\phi_m, \beta_0, ..., \beta_n, \sigma_\epsilon^2)$

Example 1 Consider the following (restricted) VAR(1) model

$$\begin{aligned} y_t &= \alpha_{11} y_{t-1} + \alpha_{12} x_{t-1} + \varepsilon_{1t}, \\ x_t &= \alpha_{22} x_{t-1} + \varepsilon_{2t}, \end{aligned}$$

where $(\varepsilon_{1t}, \varepsilon_{2t})'$ are i.i.d. $N_2(0, \Sigma)$ innovations w.r.t. (Y_{t-1}, X_{t-1}). Hence, $(y_t | Y_{t-1}, X_t; \theta)$ has a normal distribution with parameters

$$\begin{aligned} \mathsf{E}(y_t | Y_{t-1}, X_t; \theta_1) &= \alpha_{11} y_{t-1} + \alpha_{12} x_{t-1} + \frac{\sigma_{12}}{\sigma_{22}} \varepsilon_{2t}, \\ \mathsf{Var}(y_t | Y_{t-1}, X_t; \theta_1) &= \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}. \end{aligned}$$

This leads to the ADL(1,1) model

$$y_t = \phi y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t, \tag{1}$$

where $\phi = \alpha_{11}$, $\beta_0 = \frac{\sigma_{12}}{\sigma_{22}}$, $\beta_1 = \alpha_{12} - \frac{\sigma_{12}}{\sigma_{22}} \alpha_{22}$, $\mathsf{Var}(\epsilon_t) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}$, and $\theta_2 = (\alpha_{22}, \sigma_{22})'$ are variation free. Since $\mathsf{Cov}(\epsilon_t, \varepsilon_{2t}) = 0 \Rightarrow x$ is weakly exogenous for the parameters of model (1).

Granger Causality

Intuition: x causes y if the past of x has some additional predictive power for y w.r.t. the past of y itself.

Definition: x (Granger-)causes y ($x \stackrel{G}{\rightarrow} y$) if

$$\text{MSE}(y_t | Y_{t-1}, X_{t-1}) < \text{MSE}(y_t | Y_{t-1})$$

Implication: Given the following VAR(p) model for $(y_t, x_t)'$:

$$\begin{bmatrix} \alpha_{11}(L) & \alpha_{12}(L) \\ \alpha_{21}(L) & \alpha_{22}(L) \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

where $\alpha_{12}(0) = \alpha_{21}(0) = 0$, then $x \stackrel{G}{\rightarrow} y$ if $\alpha_{12}(L) \neq 0$.

Testing for Granger-causality: Identify and estimate the model

$$\alpha_{11}(L)y_t = -\alpha_{12}(L)x_t + \varepsilon_{1t}$$

Then perform a LR ratio test for the null hypothesis that x does not Granger-cause y ($x \overset{G}{\nrightarrow} y$), i.e.

$$\mathsf{H}_0 : \alpha_{12}(L) = 0$$

Example 1 (cont'd): Consider again the following bivariate model

$$\begin{aligned} y_t &= \alpha_{11} y_{t-1} + \alpha_{12} x_{t-1} + \varepsilon_{1t}, \\ x_t &= \alpha_{22} x_{t-1} + \varepsilon_{2t}. \end{aligned}$$

Then $x \overset{G}{\rightarrow} y$ and $y \overset{G}{\nrightarrow} x$.

Strong Exogeneity

Intuition: No information loss when predicting by means of a single-equation model with respect to a multivariate model. For Gaussian r.v.'s this is equivalent to require that x is weakly exogenous for the parameters of the conditional model, and $y \overset{G}{\nrightarrow} x$.

Definition: The variable x is strongly exogenous for the parameters θ_1 if the following conditions hold:

1. θ_1 and θ_2 are variation free;

2. $\mathsf{P}(x_t|Y_{t-1}, X_{t-1}; \theta_2) = \mathsf{P}(x_t|X_{t-1}; \theta_2)$.

Interpreting the ADL model

Given the ADL(m,n)

$$\phi(L)y_t = \delta + \beta(L)x_t + \epsilon_t,$$

if the roots of $\phi(L)$ are outside the unite circle, we can write

$$y_t = \phi^{-1}(L)[\delta + \beta(L)x_t + \epsilon_t] \equiv \varphi + \pi(L)x_t + \phi^{-1}(L)\epsilon_t,$$

where $\varphi = \phi^{-1}(1)\delta$, $\pi(L) = \sum_{j=0}^{\infty} \pi_j L^j$, and

$$\sum_{j=0}^{\infty} \left|\pi_j\right| < \infty,$$

which implies $\lim_{j\to\infty} \pi_j = 0$.

Dynamic impact of x_t on y_{t+h}:

$$\frac{\partial y_{t+h}}{\partial x_t} = \pi_h$$

Cumulative dynamic impact of x_t on y_{t+h}:

$$\sum_{j=0}^{h} \frac{\partial y_{t+j}}{\partial x_t} = \sum_{j=0}^{h} \pi_j$$

Long run impact of x_t on y_{t+h}:

$$\lim_{h \to \infty} \sum_{j=0}^{h} \frac{\partial y_{t+j}}{\partial x_t} = \pi(1) = \frac{\beta(1)}{\phi(1)}$$

Definition: The static long run equation is computed by putting $x_t = x^*$ and $\epsilon_t = 0$ for all t. The "equilibrium" relationship is then obtained as

$$y^* = \varphi + \pi(1)x^*,$$

or equivalently

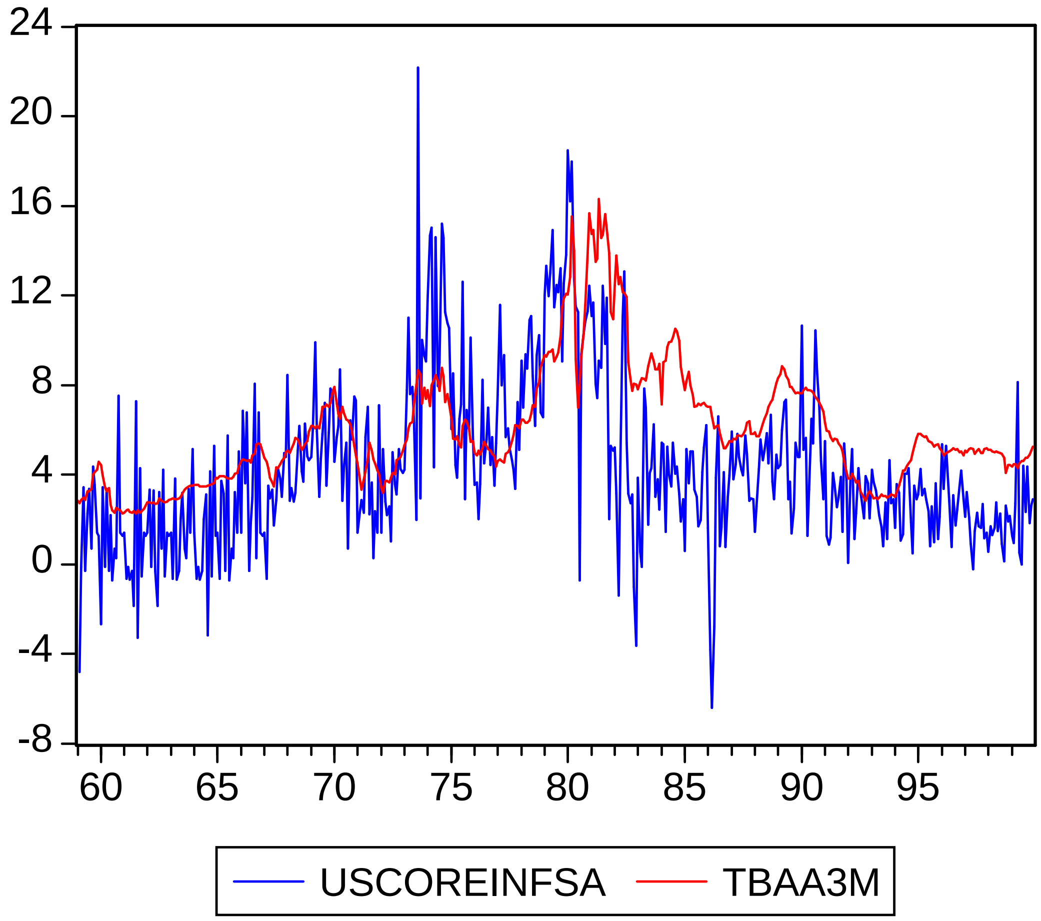
$$y^* = \frac{\delta}{\phi(1)} + \frac{\beta(1)}{\phi(1)}x^*$$

Definition: if both variables y and x are in logs, then $\pi(1)$ is the long run elasticity. It is often of interest to test if it's equal to one.

Empirical example: US Core inflation on 3-month Treasury bill, the sample is 1959.1-1999.12 (from Heij *et al.*, 2004).

ARDL example US inflation, Interest rates

US Core inflation (SA), 3-month T-bill, 59.1-99.12



ARDL(6,4) example ’reduced form’ PcGive 10

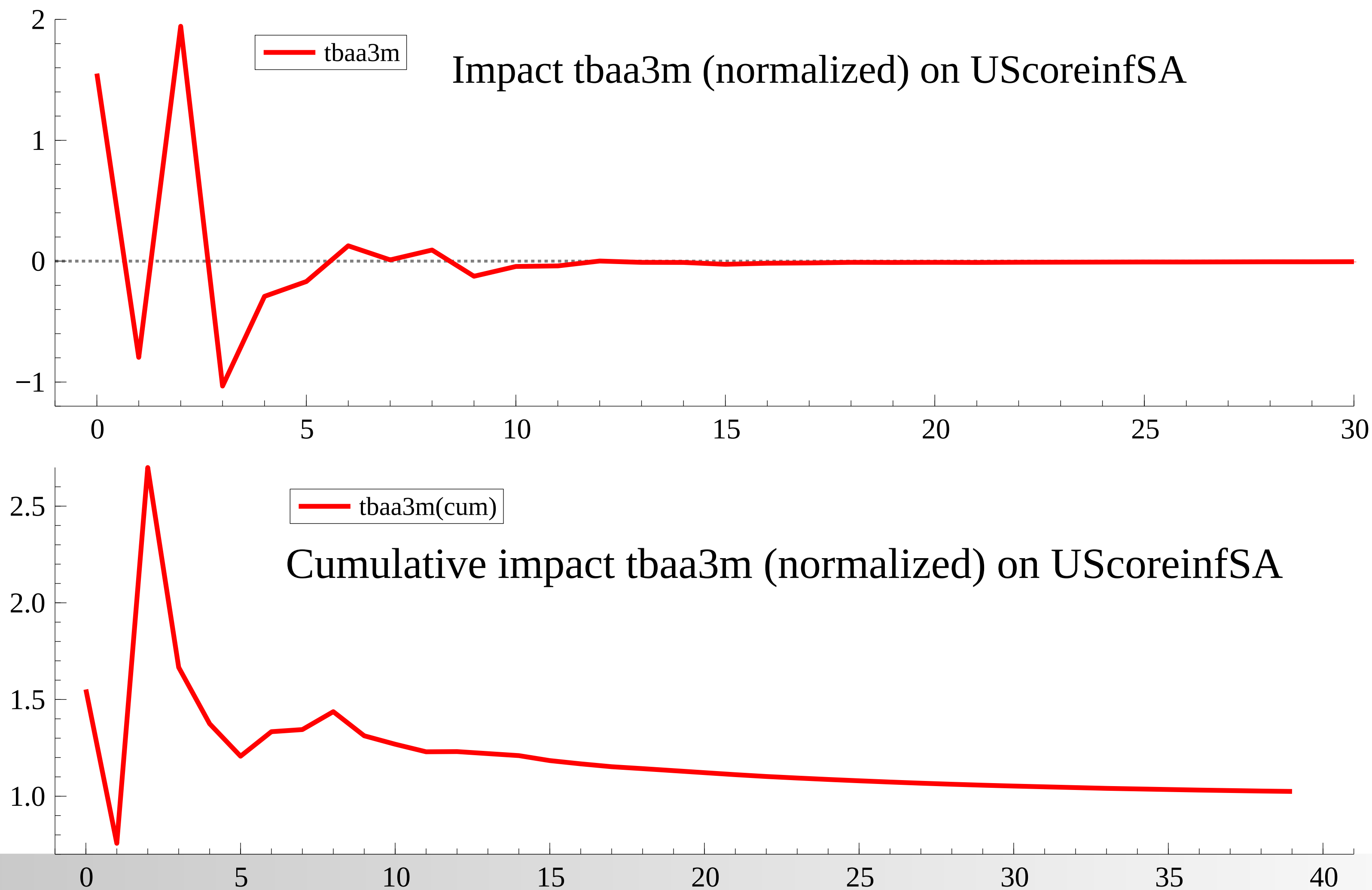
```
 EQ(11) Modelling UScoreinfSA by OLS
       The estimation sample is: 62 (1) to 99 (12)
                     Coefficient  Std.Error  t-value  t-prob
UScoreinfSA_1           0.233366    0.04750     4.91   0.000
...
UScoreinfSA_6           0.102451    0.04612     2.22   0.027
Constant                0.239006     0.3136    0.762   0.446
tbaa3m                  0.867091     0.2383     3.64   0.000
tbaa3m_1               -0.646922     0.3798    -1.70   0.089
...
tbaa3m_4               -0.340460     0.2405    -1.42   0.158
sigma                    2.54947  RSS                2885.89892
R^2                     0.542968  F(11,444) =   47.95 [0.000]**
log-likelihood          -1067.72  DW                       2.04
no. of observations          456  no. of parameters          12
mean(UScoreinfSA)        4.54782   var(UScoreinfSA)     13.8475
```

ARDL 'structural' estimation results

```
Solved static long run equation for UScoreinfSA
                  Coefficient  Std.Error  t-value  t-prob
Constant              1.07147      1.445    0.742   0.459
tbaa3m               0.558555     0.2175     2.57   0.011
ECM = UScoreinfSA - 1.07147 - 0.558555*tbaa3m;
WALD test: Chi^2(1) = 6.59755 [0.0102] *
Roots of UScoreinfSA lag polynomial:
         real         imag      modulus
      0.92014      0.00000      0.92014
      0.32600      0.61182      0.69326
      ...
     -0.34068     -0.48614      0.59363
Roots of tbaa3m lag polynomial:
         real         imag      modulus
     0.047709       1.1703       1.1713
     ..
     -0.30081      0.00000      0.30081
```

ARDL dynamic (cumulative) impact of interest

Dynamic impact interest rate on inflation (sum=1)



Error (Equilibrium) Correction Model [ECM]

Intuition: The ADL model is re-parameterized such that it is apparent as changes Δy_t react to lagged equilibrium errors $z_{t-1} \equiv y_{t-1} - \varphi - \pi(1)x_{t-1}$.

Definition: The ECM is expressed as

$$\phi^*(L)\Delta y_t = -\phi(1)z_{t-1} + \beta^*(L)\Delta x_t + \epsilon_t,$$

which is obtained from the ADL model by expanding the lag polynomials $\phi(L)$ and $\beta(L)$ as follows

$$\begin{aligned}\phi(L) &= \Delta\phi^*(L) + \phi(1)L,\\ \beta(L) &= \Delta\beta^*(L) + \beta(1)L,\end{aligned}$$

Interpretation: When y_t is higher than equilibrium value (positive z_t), y_{t+1} will adjust downwards in order to get back to equilibrium.

Example 2 Consider the cointegrated VAR(1) model

$$\Delta y_t = \alpha_1(y_{t-1} - \pi(1)x_{t-1}) + \varepsilon_{1t},$$
$$\Delta x_t = \alpha_2(y_{t-1} - \pi(1)x_{t-1}) + \varepsilon_{2t},$$

The conditional distribution $(y_t|Y_{t-1}, X_t; \theta)$ has parameters

$$\mathsf{E}(y_t|Y_{t-1}, X_t; \theta_1) = \alpha_1(y_{t-1} - \pi(1)x_{t-1}) + \frac{\sigma_{12}}{\sigma_{22}}\varepsilon_{2t},$$
$$\mathsf{Var}(y_t|Y_{t-1}, X_t; \theta_1) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}.$$

This leads to the (conditional) ECM model

$$\Delta y_t = \beta_0 \Delta x_t + \alpha(y_{t-1} - \pi(1)x_{t-1}) + \epsilon_t,$$

where $\theta_1 = (\beta_0 = \frac{\sigma_{12}}{\sigma_{22}}, \alpha = \alpha_1 - \frac{\sigma_{12}}{\sigma_{22}}\alpha_2, \pi(1), \mathsf{Var}(\epsilon_t) = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}})'$ and $\theta_2 = (\alpha_2, \pi(1), \sigma_{22})'$ are clearly not variation free. However, when $\alpha_2 = 0 \Rightarrow x$ is weakly exogenous for θ_1.

Common Factors (Roots) Analysis

Intuition: The polynomials $\phi(L)$ and $\beta(L)$ share some common roots, thus leading to a regression model with AR errors.

Definition: The ADL(1,1) model

$$(1 - \phi L)y_t = (\beta_0 + \beta_1 L)x_t + \epsilon_t \tag{2}$$

has a common factor (root) if $\beta_1/\beta_0 = -\phi$. In this case, model (2) can be rewritten as a static regression model with AR(1) errors

$$\begin{aligned} y_t &= \beta_0 x_t + \xi_t, \\ (1 - \phi L)\xi_t &= \epsilon_t. \end{aligned}$$

Definition (cont'd): The ADL(m, n) model

$$\phi(L)y_t = \beta(L)x_t + \epsilon_t \tag{3}$$

has s ($s \leq \min\{m, n\}$) common roots if

$$\begin{aligned} \phi(L) &= \alpha(L)\phi^\dagger(L), \\ \beta(L) &= \alpha(L)\beta^\dagger(L), \end{aligned}$$

where $\alpha(L)$ is a polynomial of order s. In this case, model (3) can be rewritten as an ADL($m - s, n - s$) model with AR(s) errors

$$\begin{aligned} \phi^\dagger(L)y_t &= \beta^\dagger(L)x_t + \xi_t, \\ \alpha(L)\xi_t &= \epsilon_t. \end{aligned}$$

Testing for common roots: Not easy for the general case. It can be done by the COMFAC procedure, which is a sequence Wald tests for proper non-linear cross restrictions on the coefficient polynomials $\phi(L)$ and $\beta(L)$.