ECON 4160: Econometrics–Modelling and Systems Estimation:

Computer Class # 3

André K. Anundsen

Norges Bank

September 25, 2014

Outline

Some Monte Carlo simulations (the stuff we couldn't do last time)

Data sets and batch files

Parameter stability

Models of the VAR

Single equation typology: special cases of ARDL (if time)

Weak exogeneity (if time)

What is Monte Carlo simulation? I

Say that the process that has generated the data (the data generating process, or the DGP) takes the following form:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \tag{1}$$

where ε_t is WN. Now, assume that we had some data for the period $t = 1, \ldots, T$ on y_t and x_t, and that we want to pin down β_0 and β_1 (our parameters of interest)

As long as $Cov(x_t, \varepsilon_t) = 0$, you know that the OLS estimator is BLUE! How could we confirm this by simulation?

What is Monte Carlo simulation? II

1. Fix β_0 and β_1 in (1) at some values, e.g. $\beta_0 = 2$ and $\beta_1 = 1.5$

2. Generate some numbers for the time series x_t on a sample $t = 1 \ldots, T$

3. Assume e.g. that $\varepsilon_t \sim N(0, 1)$, and draw T numbers from the standard normal distribution

4. Then, the time series y_t for the sample $t = 1, \ldots T$ will follow by definition from the DGP!

5. Estimate an equation of the form (1) by OLS and collect your β_0 and β_1 estimates; call them $\hat{\beta}_0^1$ and $\hat{\beta}_1^1$

6. Now, repeat the steps 1–5 M times, and calculate $\beta_0^{MC} = \frac{\sum_{m=1}^{M} \hat{\beta}_0^m}{M}$ and $\beta_1^{MC} = \frac{\sum_{m=1}^{M} \hat{\beta}_1^m}{M}$!

What is Monte Carlo simulation? III

As we have seen, β_0^{MC} and β_1^{MC} are nothing but the mean estimates obtained by repeating 1–5 M times. From the law of large numbers, we know that β_0^{MC} and β_1^{MC} converge to $E(\hat{\beta}_0)$ and $E(\hat{\beta}_1)$ as $M \rightarrow \infty$.

But then, we know that a measure of the biases of these estimates would be: $\beta_0^{MC} - \beta_0$ and $\beta_1^{MC} - \beta_1$. What we can also do is to consider a recursive Monte Carlo, where we e.g. vary the sample size from say $T = 20$ to $T = 500$ in increments of 20 to check the bias of the OLS estimator for different samples!

The three experiments from last time:

1. The DGP is:

$$y_t = 2 + 1.5x_t + \varepsilon_t$$

and the estimated equation is:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

2. The DGP is:

$$y_t = 2 + 1.5x_t - 2z_t + u_t$$

and the estimated equation is:

$$y_t = \gamma_0 + \gamma_1 x_t + u_t$$

3. The DGP is:

$$y_t = 2 + 0.6y_{t-1} + v_t,$$

and the estimated equation is:

$$y_t = \phi_0 + \phi_1 x_t + v_t$$

In each of the cases, what are the properties of the OLS estimator?

Data sets for today:

- ▶ KonsDataSim.zip (same as we used at CC #1 and CC #2)
- ▶ KonsDataSim2.zip
- ▶ konsdata2nor_old.zip (from Seminar # 3)

Batch (!) files for today:

- ▶ Conditional consumption equation.fl

Recursive stability

Consider the model:

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \ t = 1, 2, \ldots, T, \tag{2}$$

Last time, we looked at recursive graphs of the parameters to investigate parameter stability. We shall now consider some formal tests for structural breaks

Chow-tests of parameter stability

- ▶ If we have a hypothesis about when a structural break occurs, we can test that hypothesis
- ▶ Let T_1 denote the last period with the "old" regime and let $T_1 + 1$ denote the first period of the "new" regime:

$$y_t = \beta_1 + \beta_2 x_t + \varepsilon_t, \ t = 1, 2, \ldots, T_1 \text{ and}$$
$$y_s = \gamma_1 + \gamma_2 x_s + \epsilon_s, \ s = T_1 + 1, 2, \ldots, T.$$

  then

$$H_0: \beta_1 = \gamma_1, \ \beta_2 = \gamma_2 \text{ vs } H_1: \beta_1 \neq \gamma_1, \ \beta_2 \neq \gamma_2.$$

- ▶ There are two well known statistics for these cases, both due to Chow (1960) and referred to as *Chow tests*.

2-sample Chow-test

Let SSE be the Sum of Squared Errors (which of course corresponds to RSS or SSR):

$\diamond$ SSE_1 is for the first sample ($t = 1, 2, .., T_1$)

$\diamond$ SSE_2 is for the second sample ($t = T_1 + 1, 2, .., T_1$)

$\diamond$ $SSE_U = SSE_1 + SSE_2$

$\diamond$ SSE_R is the SSE when the whole sample is used, i.e. under H_0

The test statistic in our example is:

$$F_{Chow2} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{(T-4)}{2} \sim F(2, T-4).$$

In general, it is:

$$F_{Chow2} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T-2K}{K} \sim F(K, T-2K).$$

where K is the number of parameters

Insufficient observations Chow-test

▶ Consider $T - T_1 < K$ (the sample for the second period), where we remember that K is the number of parameters in the model. In this case we have:
  ◇ The same SSE_R as for the 2-sample Chow-test (i.e. based on the full sample)
  ◇ But SSE_U is only based on the first T_1 observations

This "predictive" Chow-test is given as

$$F_{ChowP} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{T - T_1} \sim F(T - T_1, T_1 - K)$$

Insufficient observations Chow-test

▶ If we have no clear idea about the dating of a regime shift, a graph
  with the whole sequence of predictive Chow tests is useful

▶ Chow tests rely on constant and equal variances of the disturbances.
  Hence, it is good practice to plot the recursively estimated $\hat{\sigma}$

Recursive Chow-test: 1-step test

The Test-Menu Graphics in PcGive contains three version of the F_{ChowP} test!

- ▶ *1-step test*: Both T_1 and T is increased by one observation sequentially, so that $N = T - T_1$ is always 1. Naturally, the test statistic is:

  - ▶

$$F_{ChowP} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{T - T_1}$$
$$= \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{1} \sim F(1, T_1 - K)$$

Example:

1. Estimate (2) on the sample $t = 1, ..., 10$ and $t = 1, ..., 11$ and calculate F_{ChowP}
2. Estimate (2) on the sample $t = 1, ..., 11$ and $t = 1, ..., 12$ and calculate F_{ChowP}
3. Continue like this! The graph plots the sequence of F_{ChowP}

Recursive Chow-test: Break-point Chow test

▶ *Break-point Chow test*, called *N-dn CHOWS* (N down Chow Statistic) in PcGive Here T_1 is increased by one observation, then by 2 and so on until $T_1 = T$ (T is always fixed!). Hence, $N = T - T_1$ is decreasing (N down) through the sequence of tests. The test statistic is:

$$F_{ChowP} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{T - T_1} \sim F(T - T_1, T_1 - K)$$

Example

1. Estimate (2) on the sample $t = 1, ..., 10$ and $t = 1, ..., 200$ and calculate F_{ChowP}
2. Estimate (2) on the sample $t = 1, ..., 11$ and $t = 1, ..., 200$ and calculate F_{ChowP}
3. Continue like this! The graph plots the sequence of F_{ChowP}

Recursive Chow-test: Forecast Chow test

The final version is the:

▶ *Forecast Chow test*, called *N-up CHOWS* (N up Statistic): Here T_1 is kept fixed and $T - T_1$ is first 1, then 2, and so on. Hence, $N = T - T_1$ is increasing through the sequence of tests (N up). The test statistic is:

$$F_{ChowP} = \frac{SSE_R - SSE_U}{SSE_U} \cdot \frac{T_1 - K}{T - T_1} \sim F(T - T_1, T_1 - K)$$

Example

1. Estimate (2) on the sample $t = 1, ..., 10$ and $t = 1, ..., 199$ and calculate F_{ChowP}
2. Estimate (2) on the sample $t = 1, ..., 10$ and $t = 1, ..., 200$ and calculate F_{ChowP}
3. Continue like this! The graph plots the sequence of F_{ChowP}

▶ All the Recursive Chows are scaled by (1-off) critical values. A 1% significance level is the default, so that the critical values becomes a straight line at unity

You can easily calculate the different recursive Chow test statistics yourself "by hand" by using OLS, and then compare what you get to what the program tells you (choose *Write results instead of graphing*). This is a nice way to become familiar with the tests!

Open *KonsDataSim.in*7, and we shall see how to perform these tests in PcGive!

A conditional model of the VAR

As economists we will typically be interested in building econometric *models* of the VAR *system*

Today, we will consider some alternative models of the VAR!

From VAR to ARDL I

Consider a bi-variate VAR model of first order, i.e.:

$$
\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \quad (3)
$$

A more compact notation gives:

$$
\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \quad (4)
$$

Let us now assume that $\boldsymbol{\varepsilon}_t \sim MVN(\mathbf{0}_{2\times 1}, \boldsymbol{\Sigma})$, i.e.:

$$
\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim MVN\left(\mathbf{0}_{2\times 1}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right) \quad (5)
$$

From VAR to ARDL II

It follows from (3) and (5) that $\mathbf{y}_t|\mathbf{y}_{t-1} \sim MVN\left(E(\mathbf{y}_t|\mathbf{y}_{t-1}), \mathbf{\Sigma}\right)$, where it is easy to show that:

$$
\begin{aligned}
E(\mathbf{y}_t|\mathbf{y}_{t-1}) &= E\left(\left.\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}\right| y_{1,t-1}, y_{2,t-1}\right) \\
&= \begin{pmatrix} \mu_1 + a_{11}y_{1,t-1} + a_{12}y_{2,t-1} \\ \mu_2 + a_{21}y_{1,t-1} + a_{22}y_{2,t-1} \end{pmatrix}
\end{aligned} \tag{6}
$$

From VAR to ARDL III

Since $\mathbf{y}_t|\mathbf{y}_{t-1} \sim MVN\left(E(\mathbf{y}_t|\mathbf{y}_{t-1}), \boldsymbol{\Sigma}\right)$, we know from the properties of the MVN distribution that:

$$
y_{1,t}|y_{2,t} \sim N\left(\underbrace{E(y_{1,t}|y_{1,t-1},y_{2,t-1}) - \rho_{12}\frac{\sigma_1}{\sigma_2}\left(E(y_{2,t}|y_{1,t-1},y_{2,t-1}) - y_{2,t}\right)}_{=E(y_{1,t}|y_{2,t},y_{1,t-1},y_{2,t-1})}\right.
$$

$$
\left. , \underbrace{(1-\rho_{12}^2)\sigma_1^2}_{=Var(y_{1,t}|y_{2,t},y_{1,t-1},y_{2,t-1})}\right) \tag{7}
$$

with $\rho_{12} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$

From VAR to ARDL IV

Using the expressions for $E(\mathbf{y}_t|\mathbf{y}_{t-1})$ as derived in (6) in combination
with the expression for $E(y_{1,t}|y_{2,t}, y_{1,t-1}, y_{2,t-1})$ in (7), we get that:

$$
\begin{aligned}
E(y_{1,t}|y_{2,t}, y_{1,t-1}, y_{2,t-1}) = & \underbrace{\mu_1 + a_{11}y_{1,t-1} + a_{12}y_{2,t-1}}_{E(y_{1,t}|y_{1,t-1},y_{2,t-1})} \\
& - \rho_{12}\frac{\sigma_1}{\sigma_2}\left(\underbrace{\mu_2 + a_{21}y_{1,t-1} + a_{22}y_{2,t-1}}_{E(y_{2,t}|y_{1,t-1},y_{2,t-1})} - y_{2,t} \right)
\end{aligned}
\tag{8}
$$

just collecting some terms, we get:

$$
E(y_{1,t}|y_{2,t}, \mathbf{y}_{t-1}) = \underbrace{\phi_0}_{=\mu_1 - \frac{\sigma_{12}}{\sigma_2^2}\mu_2} + \underbrace{\beta_0}_{=\frac{\sigma_{12}}{\sigma_2^2}}y_{2,t} + \underbrace{\phi_1}_{=a_{11}-\frac{\sigma_{12}}{\sigma_2^2}a_{21}}y_{1,t-1} + \underbrace{\beta_1}_{=a_{12}-\frac{\sigma_{12}}{\sigma_2^2}a_{22}}y_{2,t-1} \tag{9}
$$

From VAR to ARDL V

Define the stochastic variable ϵ_t in the following way:

$$\epsilon_t = y_{1,t} - E(y_{1,t}|y_{2,t}, \mathbf{y}_{t-1}) \tag{10}$$

Combining (9) and (10) yields

$$y_{1,t} = \phi_0 + \beta_0 y_{2,t} + \phi_1 y_{1,t-1} + \beta_1 y_{2,t-1} + \epsilon_t \tag{11}$$

which is nothing but a conditional model for $y_{1,t}$, and now you see how this may be derived from the VAR

The ARDL disturbance I

Consider the expression for the ARDL disturbance:

$$\epsilon_t = y_{1,t} - E(y_{1,t}|y_{2,t}, \mathbf{y}_{t-1})$$

Substitute in for $E(y_{1,t}|y_{2,t}, \mathbf{y}_{t-1})$ from (8) to get (after some simple re-arrangements):

$$\epsilon_t = (y_{1,t} - \mu_1 - a_{11}y_{1,t-1} - a_{12}y_{2,t-1})$$
$$- \rho_{12}\frac{\sigma_1}{\sigma_2}(y_{2,t} - \mu_2 - a_{21}y_{1,t-1} - a_{22}y_{2,t-1})$$

If you have a quick look at where we started out (equation (3)), you will recognize that the terms in the parentheses are nothing but the two VAR disturbances! Hence:

$$\epsilon_t = \varepsilon_{1,t} - \rho_{12}\frac{\sigma_1}{\sigma_2}\varepsilon_{2,t} \tag{12}$$

The ARDL disturbance II

The expression in (12) can be used to show:

$$
\begin{aligned}
E(\epsilon_t) &= 0, \ E(\epsilon_t \varepsilon_{2,t}) = 0 \\
Var(\epsilon_t) &= Var(\varepsilon_{1,t}) + Var(\rho_{12}\frac{\sigma_1}{\sigma_2}\varepsilon_{2,t}) - 2Cov\left(\varepsilon_{1,t}, \rho_{12}\frac{\sigma_1}{\sigma_2}\varepsilon_{2,t}\right) \\
&= \sigma_1^2 + \rho_{12}^2\sigma_1^2 - 2\rho_{12}\frac{\sigma_1}{\sigma_2}\sigma_{12} \\
&= \sigma_1^2 + \rho_{12}^2\sigma_1^2 - 2\rho_{12}\frac{\sigma_1}{\sigma_2}\underbrace{\rho_{12}\sigma_1\sigma_2}_{=\sigma_{12}} = \sigma_1^2(1 - \rho_{12}) \\
E(y_{2,t}\epsilon_t) &= 0 \ \forall \ \text{(for all)} \ t
\end{aligned}
$$

The statistical system described by (3) and (5) can now be expressed in
model form:

$$y_{1,t} = \phi_0 + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \phi_1 y_{1,t-1} + \epsilon_t \tag{13}$$
$$y_{2,t} = \mu_2 + a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + \varepsilon_{2,t} \tag{14}$$

$$E(\epsilon_t) = 0 \ \forall \ t$$
$$Var(\epsilon_t) = \sigma_1^2 (1 - \rho_{12}) \ \forall \ t$$
$$E(y_{2,t}\epsilon_t) = 0 \ \forall \ t$$
$$E(\varepsilon_{2,t}\epsilon_t) = 0 \ \forall \ t$$

Regressions models I

1. When estimating a linear regression model, we are estimating a conditional expectation that is derived from a system of equations, e.g. a simple bivariate VAR
   - We are therefore estimating a *partial* system!
2. The full econometric model of the system consists of the *conditional model* (13), the *marginal model* (14), and the disturbances ϵ_t and $\varepsilon_{2,t}$
3. When $Cov(\epsilon_t, y_{2,t}) = 0$, $y_{2,t}$ is *exogenous* in the conditional model
4. OLS estimation is efficient for Gaussian (i.e., normal) disturbances and gives the Maximum Likelihood estimators for ϕ_0, β_0, ϕ_1 and β_1

Regressions models II

5. This means that there is no information in the marginal model that can help us improve on the estimates of ϕ_0, β_0, ϕ_1 and β_1 that we get from the conditional model

6. We say that $y_{2,t}$ is a *weakly exogenous* variable for the *parameters of interest*. In our case: ϕ_0, β_0, ϕ_1 and β_1 and $Var[\epsilon_t]$

Let's see if we can confirm the results we just derived! I

Load in the in7 file contained in the **KonsDataSim2.zip** (on the web).
Now, let us do the following:

1. Back rows: Estimate a bi-variate VAR of first order in consumption
   (C) and income (I), where $y_{1,t} = C_t$ and $y_{2,t} = I_t$
2. First rows: Estimate an $ARDL(1,1)$ model for a conditional
   consumption equation, i.e. $C_t | I_t, C_{t-1}, I_{t-1}$

Both groups use the full sample (1959–2007)!

Let's see if we can confirm the results we just derived! II

I have cheated, and calculated the empirical variance-covariance matrix
of the VAR residuals to save some time. It is given by:

$$\begin{pmatrix} 10290.0736 & 4655.303596 \\ 4655.303596 & 4310.579025 \end{pmatrix}$$

Can we by combining the VAR estimates obtained by the guys at the
back rows and the cov-matrix above guess what results the guys at the
front row got?

Let's see if we can confirm the results we just derived! III

Again, I have cheated and calculated in advance, but I guess the front rows got the following:

$$\hat{\phi}_0 = \hat{\mu}_1 - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_2^2}\hat{\mu}_2 = 72.0987 - \frac{4655.303596}{4310.579025}22.0776 = 48.25551557$$

$$\hat{\beta}_0 = \frac{\hat{\sigma}_{12}}{\hat{\sigma}_2^2} = \frac{4655.303596}{4310.579025} = 1.079971755$$

$$\hat{\phi}_1 = \hat{a}_{11} - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_2^2}\hat{a}_{21} = 0.702311 - \frac{4655.303596}{4310.579025}0.120492 = 0.572183043$$

$$\hat{\beta}_1 = \hat{a}_{21} - \frac{\hat{\sigma}_{12}}{\hat{\sigma}_2^2}\hat{a}_{22} = 0.351689 - \frac{4655.303596}{4310.579025}0.784072 = -0.49508661$$

Generalizations

- ▶ The results 1-3 above do – in general – not depend on normality (it is just a simplification)
- ▶ In particular: Normality of $y_{2,t}$ is *not* required for $E[\epsilon_t] = 0$ and $E[\epsilon_t y_{2,t}] = 0$
- ▶ The results $E[\epsilon_t] = 0$ and $E[\epsilon_t y_{2,t}] = 0$ do not depend on linearity. More generally, we have

$$y_{1,t} = E[y_{1,t} \mid y_{2,t}, y_{1,t-1}, y_{2,t-1}] + \epsilon_t$$

with $E[\epsilon_t] = 0$ and $E[\epsilon_t y_{2,t}] = 0$ for a *non-linear* conditional expectation function $E[y_{1,t} \mid y_{2,t}, y_{1,t-1}, y_{2,t-1}]$

- ▶ Generalizations from one to k explanatory variables and p lags are straight-forward. We get:

$$y_{1,t} = E[y_{1,t} \mid y_{1,t-1}, \ldots, y_{1,t-p}, y_{2,t}, \ldots, y_{2,t-p}, \ldots, y_{k+1,t}, \ldots y_{k+1,t-p}] + \epsilon_t$$

and the linear multiple regression is a special case.

The ecm as a model of the VAR

In a similar way as for the VAR(1) model, we can represent a VAR(2) on **model form**:

$$
\begin{aligned}
y_{1,t} &= \phi_0 + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \phi_1 y_{1,t-1} \\
&\quad + \beta_2 y_{2,t-2} + \phi_2 y_{1,t-2} + \epsilon_t
\end{aligned}
\tag{15}
$$

$$
\begin{aligned}
y_{2,t} &= \mu_2 + a_{21,1} y_{1,t-1} + a_{22,1} y_{2,t-1} \\
&\quad + a_{21,2} y_{1,t-2} + a_{22,2} y_{2,t-2} + \varepsilon_{2,t}
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
E(\epsilon_t) &= 0 \ \forall \ t \\
Var(\epsilon_t) &= \sigma_1^2 \left(1 - \rho_{12}\right) \ \forall \ t \\
E(y_{2,t} \epsilon_t) &= 0 \ \forall \ t \\
E(\varepsilon_{2,t} \epsilon_t) &= 0 \ \forall \ t
\end{aligned}
$$

The equilibrium correction model (ECM) as a reparameterization of the ARDL model

The conditional model for $y_{1,t}$ in (15) is called an ARDL(2,2) model:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

The equilibrium correction model (ECM) as a reparameterization of the ARDL model

The conditional model for $y_{1,t}$ in (15) is called an ARDL(2,2) model:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

Subtract and add $\phi_2 y_{1,t-1}$ and $\beta_2 y_{2,t-1}$ on the RHS. We then get:

The equilibrium correction model (ECM) as a reparameterization of the ARDL model

The conditional model for $y_{1,t}$ in (15) is called an ARDL(2,2) model:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

Subtract and add $\phi_2 y_{1,t-1}$ and $\beta_2 y_{2,t-1}$ on the RHS. We then get:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2}$$
$$+ \phi_2 y_{1,t-1} - \phi_2 y_{1,t-1} + \beta_2 y_{2,t-1} - \beta_2 y_{2,t-1} + \epsilon_t$$

The equilibrium correction model (ECM) as a reparameterization of the ARDL model

The conditional model for $y_{1,t}$ in (15) is called an ARDL(2,2) model:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

Subtract and add $\phi_2 y_{1,t-1}$ and $\beta_2 y_{2,t-1}$ on the RHS. We then get:

$$\begin{aligned}
y_{1,t} = {} & \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} \\
& + \phi_2 y_{1,t-1} - \phi_2 y_{1,t-1} + \beta_2 y_{2,t-1} - \beta_2 y_{2,t-1} + \epsilon_t
\end{aligned}$$

Collecting some terms, we have:

The equilibrium correction model (ECM) as a reparameterization of the ARDL model

The conditional model for $y_{1,t}$ in (15) is called an ARDL(2,2) model:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

Subtract and add $\phi_2 y_{1,t-1}$ and $\beta_2 y_{2,t-1}$ on the RHS. We then get:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2}$$
$$+ \phi_2 y_{1,t-1} - \phi_2 y_{1,t-1} + \beta_2 y_{2,t-1} - \beta_2 y_{2,t-1} + \epsilon_t$$

Collecting some terms, we have:

$$y_{1,t} = \phi_0 + (\phi_1 + \phi_2) y_{1,t-1} - \phi_2 (y_{1,t-1} - y_{1,t-2}) + \beta_0 y_{2,t}$$
$$+ (\beta_1 + \beta_2) y_{2,t-1} - \beta_2 (y_{2,t-1} - y_{2,t-2}) + \epsilon_t$$

ECM cont'd

Let $\Delta z_{t-i} == L^i z_t - L^{i+1} z_t = L^i (1 - L) z_t = z_{t-i} - z_{t-i-1}$ for any variable z (in our case y_1 and y_2). We can then write the above expression as:

$$y_{1,t} = \phi_0 + (\phi_1 + \phi_2) y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2) y_{2,t-1}$$
$$- \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

$$y_{1,t} - y_{1,t-1} = \phi_0 + (\phi_1 + \phi_2)y_{1,t-1} - \phi_2\Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2)y_{2,t-1}$$
$$- \beta_2\Delta y_{2,t-1} - y_{1,t-1} + \beta_0 y_{2,t-1} - \beta_0 y_{2,t-1} + \epsilon_t$$

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

$$y_{1,t} - y_{1,t-1} = \phi_0 + (\phi_1 + \phi_2)y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2)y_{2,t-1}$$
$$- \beta_2 \Delta y_{2,t-1} - y_{1,t-1} + \beta_0 y_{2,t-1} - \beta_0 y_{2,t-1} + \epsilon_t$$

Again, just collecting terms and using $\Delta z_{t-i} = z_{t-i} - z_{t-i-1}$, we have:

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

$$y_{1,t} - y_{1,t-1} = \phi_0 + (\phi_1 + \phi_2)y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2)y_{2,t-1}$$
$$- \beta_2 \Delta y_{2,t-1} - y_{1,t-1} + \beta_0 y_{2,t-1} - \beta_0 y_{2,t-1} + \epsilon_t$$

Again, just collecting terms and using $\Delta z_{t-i} = z_{t-i} - z_{t-i-1}$, we have:

$$\Delta y_{1,t} = \phi_0 + (\phi_1 + \phi_2 - 1)y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1}$$
$$+ (\beta_0 + \beta_1 + \beta_2)y_{2,t-1} + \epsilon_t$$

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

$$
y_{1,t} - y_{1,t-1} = \phi_0 + (\phi_1 + \phi_2)y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2)y_{2,t-1}
$$
$$
- \beta_2 \Delta y_{2,t-1} - y_{1,t-1} + \beta_0 y_{2,t-1} - \beta_0 y_{2,t-1} + \epsilon_t
$$

Again, just collecting terms and using $\Delta z_{t-i} = z_{t-i} - z_{t-i-1}$, we have:

$$
\Delta y_{1,t} = \phi_0 + (\phi_1 + \phi_2 - 1)y_{1,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1}
$$
$$
+ (\beta_0 + \beta_1 + \beta_2)y_{2,t-1} + \epsilon_t
$$

But then, we can also write this as:

ECM cont'd

Now, subtract and add $\beta_0 y_{2,t-1}$ on the RHS and also subtract $y_{1,t-1}$ on both the RHS and the LHS side. We then get:

$$y_{1,t} - y_{1,t-1} = \phi_0 + (\phi_1 + \phi_2)y_{1,t-1} - \phi_2\Delta y_{1,t-1} + \beta_0 y_{2,t} + (\beta_1 + \beta_2)y_{2,t-1}$$
$$- \beta_2\Delta y_{2,t-1} - y_{1,t-1} + \beta_0 y_{2,t-1} - \beta_0 y_{2,t-1} + \epsilon_t$$

Again, just collecting terms and using $\Delta z_{t-i} = z_{t-i} - z_{t-i-1}$, we have:

$$\Delta y_{1,t} = \phi_0 + (\phi_1 + \phi_2 - 1)y_{1,t-1} - \phi_2\Delta y_{1,t-1} + \beta_0\Delta y_{2,t} - \beta_2\Delta y_{2,t-1}$$
$$+ (\beta_0 + \beta_1 + \beta_2)y_{2,t-1} + \epsilon_t$$

But then, we can also write this as:

$$\Delta y_{1,t} = \phi_0 + (\phi_1 + \phi_2 - 1)\left(y_{1,t-1} - \frac{(\beta_0 + \beta_1 + \beta_2)}{(1 - \phi_1 + \phi_2)}y_{2,t-1}\right)$$
$$- \phi_2\Delta y_{1,t-1} + \beta_0\Delta y_{2,t} - \beta_2\Delta y_{2,t-1} + \epsilon_t$$

ECM cont'd

Now, let us define:

$$\alpha = \phi_1 + \phi_2 - 1$$
$$\gamma = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 + \phi_2} = -\frac{\beta_0 + \beta_1 + \beta_2}{\alpha}$$

ECM cont'd

Now, let us define:

$$\alpha = \phi_1 + \phi_2 - 1$$
$$\gamma = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 + \phi_2} = -\frac{\beta_0 + \beta_1 + \beta_2}{\alpha}$$

But then, we have:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - \gamma y_{2,t-1}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t \tag{17}$$

ECM cont'd

Now, let us define:

$$\alpha = \phi_1 + \phi_2 - 1$$
$$\gamma = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 + \phi_2} = -\frac{\beta_0 + \beta_1 + \beta_2}{\alpha}$$

But then, we have:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - \gamma y_{2,t-1}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t \tag{17}$$

Which is nothing but the equilibrium/error correction model!

Estimation of Long-run multipliers by NLS

Note that the ecm in (**??**) is non-linear in parameters! In this case, we could estimate the equation using NLS, which uses numerical optimization methods (See also Lecture # 3)

Estimation of long-run multipliers by OLS

An alternative is to rewrite the model into a model that is linear in parameters and use OLS:

$$\Delta y_{1,t} = \phi_0 + \alpha y_{1,t-1} + \eta y_{2,t-1} - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

where I just defined $\eta = -\alpha\gamma$. What would be your estimate of γ in this case? How would you calculate the standard error of gamma?

Calculating long-run multipliers based on the ARDL

We have shown that the mapping from the ARDL parameters to the long-run multiplier is:

$$\gamma = \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 + \phi_2} = -\frac{\beta_0 + \beta_1 + \beta_2}{\alpha}$$

with

$$\alpha = \phi_1 + \phi_2 - 1$$

Hence, we could back out the long-run multiplier from the estimated parameters of the ARDL model (fortunately, PcGive has a nice command that does the calculation (using the delta method) for us!!)

An example: The conditional consumption model from Seminar # 3

Consider the following ARDL(4,4) model for consumption in Norway:

$$
LCP_t = \mu + \sum_{i=1}^{4} a_{CC,i} LCP_{t-i} + \sum_{i=0}^{4} a_{CR,i} LRCa_{t-i} + \sum_{i=0}^{4} a_{CF,i} LF_{t-i}
$$
$$
Seasonals + \varepsilon_t
$$

An example: The conditional consumption model from Seminar # 3

Consider the following ARDL(4,4) model for consumption in Norway:

$$LCP_t = \mu + \sum_{i=1}^{4} a_{CC,i} LCP_{t-i} + \sum_{i=0}^{4} a_{CR,i} LRCa_{t-i} + \sum_{i=0}^{4} a_{CF,i} LF_{t-i}$$
$$Seasonals + \varepsilon_t$$

Ignoring the dummies for simplicity, we can rewrite the conditional consumption model as:

$$\Delta LCP_t = \mu + \sum_{i=1}^{3} \phi_i \Delta LCP_{t-i} + \sum_{i=0}^{3} \gamma_i \Delta LRCa_{t-i} + \sum_{i=0}^{3} \eta_i \Delta LF_{t-i}$$
$$+ \alpha (LCP_{t-1} - \beta_1 LRCa_{t-1} - \beta_2 LF_{t-1}) + \varepsilon_t$$

where the $\phi's$, $\gamma's$ and $\eta's$ are functions of the parameters in the conditional consumption function.

ECM cont'd

α, β_1 and β_2 are defined in the following way:

$$
\alpha = a_{CC1} + a_{CC2} + a_{CC3} + a_{CC4} - 1
$$
$$
\beta_1 = \frac{a_{CR1} + a_{CR2} + a_{CR3} + a_{CR4}}{1 - a_{CC1} + a_{CC2} + a_{CC3} + a_{CC4}} = -\frac{a_{CR1} + a_{CR2} + a_{CR3} + a_{CR4}}{\alpha}
$$
$$
\beta_2 = \frac{a_{CF1} + a_{CF2} + a_{CF3} + a_{CF4}}{1 - a_{CC1} + a_{CC2} + a_{CC3} + a_{CC4}} = -\frac{a_{CF1} + a_{CF2} + a_{CF3} + a_{CF4}}{\alpha}
$$

where we also know that:

$$\Delta LCP_t = \mu + \sum_{i=1}^{3} \phi_i \Delta LCP_{t-i} + \sum_{i=0}^{3} \gamma_i \Delta LRCa_{t-i} + \sum_{i=0}^{3} \eta_i \Delta LF_{t-i}$$
$$+ \alpha(LCP_{t-1} - \beta_1 LRCa_{t-1} - \beta_2 LF_{t-1}) + \varepsilon_t$$

can be written as:

$$\Delta LCP_t = \mu + \sum_{i=1}^{3} \phi_i \Delta LCP_{t-i} + \sum_{i=0}^{3} \gamma_i \Delta LRCa_{t-i} + \sum_{i=0}^{3} \eta_i \Delta LF_{t-i}$$
$$+ \alpha LCP_{t-1} + \psi_1 LRCa_{t-1} + \psi_2 LF_{t-1} + \varepsilon_t$$

where $\psi_1 = -\alpha\beta_1$ and $\psi_2 = -\alpha\beta_2$.

Using what we have just learnt, calculate the long-run multipliers for the conditional consumption model using the **KonsDataNor2.in7** data set on the sample 1985q1–2004q4! Estimate the ARDL and calculate the static long-run solution, estimate the restricted equilibrium correction model by NLS and the unrestricted equilibrium correction model by OLS. Show that you get exactly the same answer regarding the long-run multipliers!!! To increase efficiency, you can also use the batch file **Conditional consumption equation.fl**.

Why do we call this model an ECM?

Remember that the conditional model for $y_{1,t}$ is:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

Why do we call this model an ECM?

Remember that the conditional model for $y_{1,t}$ is:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

In a static long-run equilibrium, we have:

$$
\begin{aligned}
y_{1,t} = y_{1,t-1} = y_{1,t-2} = y_1^* \\
y_{2,t} = y_{2,t-1} = y_{2,t-2} = y_2^* \\
\epsilon_t = \epsilon^* = 0
\end{aligned}
$$

Why do we call this model an ECM?

Remember that the conditional model for $y_{1,t}$ is:

$$y_{1,t} = \phi_0 + \phi_1 y_{1,t-1} + \phi_2 y_{1,t-2} + \beta_0 y_{2,t} + \beta_1 y_{2,t-1} + \beta_2 y_{2,t-2} + \epsilon_t$$

In a static long-run equilibrium, we have:

$$y_{1,t} = y_{1,t-1} = y_{1,t-2} = y_1^*$$
$$y_{2,t} = y_{2,t-1} = y_{2,t-2} = y_2^*$$
$$\epsilon_t = \epsilon^* = 0$$

Thus, we find that:

$$y_1^* = \phi_0 + \phi_1 y_1^* + \phi_2 y_1^* + \beta_0 y_2^* + \beta_1 y_2^* + \beta_2 y_2^*$$

Solving for y_1^* , we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

Solving for y_1^*, we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

But this means that $\gamma y_2^* = y_1^* + \frac{\phi_0}{\alpha}$! Substituting this back into (17), we find:

Solving for y_1^*, we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

But this means that $\gamma y_2^* = y_1^* + \frac{\phi_0}{\alpha}$! Substituting this back into (17), we find:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - y_{1,t-1}^* - \frac{\phi_0}{\alpha}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

Solving for y_1^*, we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

But this means that $\gamma y_2^* = y_1^* + \frac{\phi_0}{\alpha}$! Substituting this back into (17), we find:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - y_{1,t-1}^* - \frac{\phi_0}{\alpha}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

which also can be written as:

Solving for y_1^*, we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

But this means that $\gamma y_2^* = y_1^* + \frac{\phi_0}{\alpha}$! Substituting this back into (17), we find:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - y_{1,t-1}^* - \frac{\phi_0}{\alpha}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

which also can be written as:

$$\Delta y_{1,t} = \phi_0 - \alpha \frac{\phi_0}{\alpha} + \alpha(y_{1,t-1} - y_{1,t-1}^*) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

Solving for y_1^*, we get:

$$y_1^* = \frac{\phi_0}{1 - \phi_1 - \phi_2} + \frac{\beta_0 + \beta_1 + \beta_2}{1 - \phi_1 - \phi_1} y_2^* = -\frac{\phi_0}{\alpha} - \frac{\beta_0 + \beta_1 + \beta_2}{\alpha} y_2^* = -\frac{\phi_0}{\alpha} + \gamma y_2^*$$

But this means that $\gamma y_2^* = y_1^* + \frac{\phi_0}{\alpha}$! Substituting this back into (17), we find:

$$\Delta y_{1,t} = \phi_0 + \alpha(y_{1,t-1} - y_{1,t-1}^* - \frac{\phi_0}{\alpha}) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

which also can be written as:

$$\Delta y_{1,t} = \phi_0 - \alpha \frac{\phi_0}{\alpha} + \alpha(y_{1,t-1} - y_{1,t-1}^*) - \phi_2 \Delta y_{1,t-1} + \beta_0 \Delta y_{2,t} - \beta_2 \Delta y_{2,t-1} + \epsilon_t$$

Thus, we see that, as long as $\alpha = \phi_1 + \phi_2 - 1 < 0 \Rightarrow \phi_1 + \phi_2 < |1|$, Δy_1 will fall (equilibrium correct) when $y_1 > y_1^*$ in the previous period!!

Remember the ARDL model in its general form:

$$y_t = \beta_0 + \phi_1 y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t \qquad (18)$$

$$\phi_1 = 0 \qquad \begin{array}{c} \text{Distributed lag (DL} \\ y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t. \end{array}$$

Growth rate (GR)

$$\begin{array}{c} \phi_1 = 1 \\ \beta_1 + \beta_2 = 0 \end{array} \qquad \Delta y_t = \beta_0 + \beta_1 \Delta x_t + \varepsilon_t.$$

Random Walk (RW)

$$\begin{array}{c} \phi_1 = 1 \\ \beta_1 = \beta_2 = 0 \end{array} \qquad y_t = \beta_0 + y_{t-1} + \varepsilon_t.$$

Static model

$$\begin{array}{c} \phi_1 = 0 \\ \beta_2 = 0 \end{array} \qquad y_t = \beta_0 + \beta_1 x_t + \varepsilon_t.$$

ECM

$$|\phi_1| < 1 \qquad \begin{array}{c} \Delta y_t = \beta_0 + \beta_1 \Delta x_t + (\phi_1 - 1) y_{t-1} \\ + (\beta_1 + \beta_2) x_{t-1} \varepsilon_t. \end{array}$$

And many more, including polynomial lag distributions and geometric lag distribution, see EB note DL which can be used to represent longer lags by a small number of parameters

▶ Note in particular that the ECM representation does not impose any restrictions on the ARDL model, except $|\phi_1| < 1$ (dynamic stability).

▶ The partial derivatives

$$\frac{\partial y_t}{\partial x_{t-j}}$$

also called *dynamic multipliers,* or l*ag-weights*, are easy to obtain in PcGive after estimation, as are the *long-run multipliers*, as we have seen

▶ As all parameters of an econometric model, also dynamic and long-run multipliers can become badly biased if a mis-specified model is estimated

Weak exogeneity of $y_{2,t}$ in the conditional model

OLS gives ML estimates of the parameters of the ARDL model

$$y_{1,t} = \phi_0 + \beta_0 y_{2,t} + \phi_1 y_{1,t-1} \beta_1 y_{2,t-1} + \epsilon_t \tag{19}$$

$y_{2,t}$ is therefore weakly exogenous in (19) despite the fact that $y_{2,t}$ is an endogenous variable in the VAR:

$$y_{1,t} = a_{11} y_{1,t-1} + a_{12} y_{2,t-1} + \varepsilon_{1,t} \tag{20}$$

$$y_{2,t} = a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + \varepsilon_{2,t} \tag{21}$$

$$\left(\begin{array}{c} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{array} \right) \sim MVN \left(\mathbf{0}_{2 \times 1}, \left(\begin{array}{cc} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{array} \right) \right) \tag{22}$$

▶ There is a difference between a variable being endogenous in a statistical system like (20)-(22) and being endogenous in a *model* of the statistical system, such as (19)

▶ $y_{2,t}$ is *weakly exogenous* for the parameters in (19) because we do not gain anything in terms of efficiency by estimating (19) jointly with the marginal equation (21).

Again, this is a consequence of the conditioning, which also gives

$$E(\epsilon_t \varepsilon_{2,t}) = 0 \Rightarrow E(\epsilon_t y_{2,t}) = 0$$

so $y_{2,t}$ is exogenous in the econometric sense that is used in most textbooks (sometimes referred to as the condition of *strict exogeneity*.)

Parameters of interest and weak exogeneity

- ▶ How helpful and relevant is the weak exogeneity of a variable in a conditional (regression) model?
- ▶ It *is* relevant if the parameters that we want to estimate, *the parameters of interest*, are the parameters of the conditional model!
- ▶ If the parameters of interest are not the conditional model, then the weak exogeneity of $y_{2,t}$ is not very helpful
- ▶ The solution is to change to a *different econometric model* of the system
- ▶ The other model is estimated by *other methods* than OLS

Predeterminedness

Consider again the two-variable ARDL model that we have derived from the bivariate VAR(1) model

$$y_{1,t} = \phi_0 + \beta_0 y_{2,t} + \phi_1 y_{1,t-1} + \beta_1 y_{2,t-1} + \epsilon_t \tag{23}$$

we have that

$$E(y_{1,t-1}\epsilon_{t+j}) = 0, \text{and } E(y_{2,t-1}\epsilon_{t+j}) = 0 \ \forall \ j > 0$$

by conditioning on history of the system, and

$$E(y_{2,t}\epsilon_{t+j}) = 0 \ \forall \ j > 0$$

by conditioning on $y_{2,t}$

Heuristically, we cannot claim strict exogeneity

$$E(y_{1,t-1}\epsilon_{t\pm j}) = 0 \ \forall \ j \qquad (24)$$

Intuitively, this is because $y_{1,t-1}$ must be correlated with $\epsilon_{t-1}, \epsilon_{t-2}$ and older disturbances through the solution of the equation for $y_{1,t}$

- ▶ (24) defines $y_{1,t-1}$ as a pre-determined variable.
- ▶ $y_{2,t}$ and $y_{2,t-1}$ are either exogenous or predetermined (depending on Granger causality, which we will discuss in more detail the next time)
- ▶ With pre-determinedness OLS estimators are biased in small samples, but they remain consistent estimators in stationary systems
- ▶ The size of the bias is seldom very large (as we have already seen), and it declines with ϕ_1