Macroeconometrics

Topic 3: Multivariate Time Series

by Vanessa Berenguer-Rico

University of Oxford

Michaelmas 2016

Multivariate Time Series

- Spurious Regressions & Co-integration
- Single Equation Dynamic Models: ARDL & ECM models

Spurious Regressions

Spurious Regressions

- Spurious regressions have a long tradition in statistics
- Yule (1926): "Why do we sometimes get nonsense-correlations between time-series?"
- *"It is fairly familiar knowledge that we sometimes obtain between quantities varying with the time (time-variables) quite high correlations to which we cannot attach any physical significance whatever, although under the ordinary test the correlation would be held to be certainly "significant""*

Spurious Regressions

- Real Data Example: Yule (1926)

- Proportion of Church of England marriages to all marriages for the years 1866-1911 inclusive and standardized mortality per 1000 persons for the same years

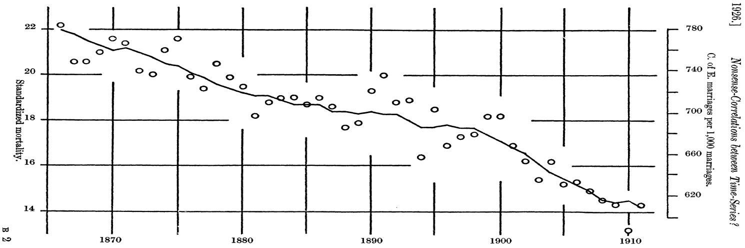


FIG. 1.—Correlation between standardized mortality per 1,000 persons in England and Wales (circles), and the proportion of Church of England marriages per 1,000 of all marriages (line), 1866–1911. $r = +0{\cdot}9512$.

- Correlation coefficient = 0.9512

Spurious Regression

- *"As the occurrence of such "nonsense-correlations" makes one mistrust the serious arguments that are sometimes put forward on the basis of correlations between time series (...) it is important to clear up the problem* ***how they arise and in what special cases****."* Yule (1926)

- *"When we find that a theoretical formula applied to a particular case gives results which common sense judges to be incorrect, it is generally as well to* ***examine the particular assumptions*** *from which it was deduced, and see which of them are inapplicable to the case in point."* Yule (1926)

Spurious Regressions

- Yule (1926): Experimental exercise

- He simulated the distribution of the empirical correlation coefficient calculated from two independent i.i.d. processes, from two independent random walks, and from two independent cumulated random walks

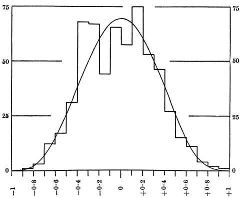


Fig. 16.—Frequency-distribution of 600 correlations between samples of 10 observations from random series (Table VIII).

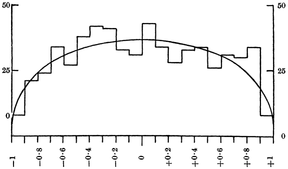


Fig. 17.—Frequency-distribution of 600 correlations between samples of 10 observations from conjunct series with random differences (Series X_1, Table IX).

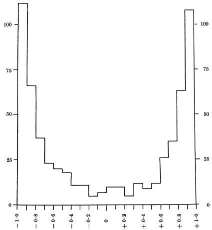


Fig. 18.—Frequency-distribution of 600 correlations between samples of 10 observations from conjunct series with conjunct differences (Table XI).

Spurious Regressions

- Granger and Newbold (1974): "Spurious Regressions in Econometrics"

- *"It is very common to see reported in applied econometric literature time series regression equations with an apparently high degree of fit, measured by the coefficient of multiple correlation R^2 or the corrected coefficient $\bar{R}^2$, but with an extremely low value for the Durbin-Watson statistic."*

- How nonsense regressions can arise? A simulation study:

- Let $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$ with $y_0 = 0$, $x_0 = 0$, and $v_t \sim i.i.d.\,(0,1)$ independent of $w_t \sim i.i.d.\,(0,1)$ and consider the least squares regression

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$$

Spurious Regressions

Let $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$ with $y_0 = 0$, $x_0 = 0$, and $v_t \sim i.i.d.\,(0,1)$ independent of $w_t \sim i.i.d.\,(0,1)$ and consider the least squares regression $y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$

A draw	$T = 99$	$T = 999$
$\hat{\alpha}$	-0102	-11.658
$t_{\hat{\alpha}}$	-0.170	-26.032
$\hat{\beta}$	0.605	0.610
$t_{\hat{\beta}}$	10.821	65.532
R^2	0.546	0.811
DW	0.151	0.032

Spurious Regressions

A Monte Carlo		percentiles			
		5%	50%	95%	
$T = 99$	$t_{\hat{\alpha}}$	-17.384	0.361	18.955	
	$t_{\hat{\beta}}$	-10.987	0.012	11.247	$rf_{1.96}$=0.756
	R^2	0.001	0.168	0.659	$m(R^2)$=0.232
	DW	0.047	0.144	0.372	$m(DW)$=0.171
$T = 999$	$t_{\hat{\alpha}}$	-56.073	-1.656	56.769	
	$t_{\hat{\beta}}$	-37.819	0.516	39.271	$rf_{1.96}$=0.927
	R^2	0.001	0.171	0.695	$m(R^2)$=0.240
	DW	0.004	0.016	0.038	$m(DW)$=0.018

Spurious Regressions

- Phillips (1986): "Understanding Spurious Regressions in Econometrics"
- *"The present paper develops an **asymptotic theory** for regressions that relate quite general integrated random processes. This includes **spurious regressions** of the Granger-Newbold type as a special case. It turns out that the correct asymptotic theory goes a long way towards explaining the experimental results that these authors obtained. In many cases their findings are quite predictable from the true asymptotic behavior of the relevant statistics."*

Spurious Regressions

- *"Thus, our theory demonstrate that in the Granger-Newbold regressions of independent random walks the usual t-ratio significance test does not possess a limiting distribution, but actually diverges as the sample size $T \to \infty$."*

- *"Inevitably, therefore, the bias in this test towards the rejection of no relationship (based on a nominal critical value of 1.96) will increase with T."*

- *"We also show that the Durbin-Watson statistic actually converges in probability to zero, while the regression R^2 has a non-degenerate limiting distribution as $T \to \infty$."*

Spurious Regressions

Let $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$ with $y_0 = 0$, $x_0 = 0$, and $v_t \sim i.i.d.\,(0,1)$ independent of $w_t \sim i.i.d.\,(0,1)$ and consider the least squares regression

$$y_t = \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$$

where

$$\hat{\beta} = \frac{\sum_1^T y_t x_t}{\sum_1^T x_t^2}$$

Spurious Regressions

(a) Denominator

$$T^{-2}\sum_{1}^{T} x_t^2 \Longrightarrow \sigma_w^2 \int_0^1 W(r)^2\, dr$$

(b) Numerator

$$T^{-2}\sum_{1}^{T} y_t x_t \Longrightarrow \sigma_v \sigma_w \int_0^1 V(r)\, W(r)\, dr$$

Spurious Regressions

Therefore,

$$\hat{\beta} = \frac{\sum_1^T y_t x_t}{\sum_1^T x_t^2} \xrightarrow{d} \frac{\sigma_v}{\sigma_w} \frac{\int_0^1 V\left(r\right) W\left(r\right) dr}{\int_0^1 W\left(r\right)^2 dr}$$

That is the OLS estimator converges to a random variable

Spurious Regressions

It can also be shown that

$$T^{-1/2} t_{\beta} \Longrightarrow D,$$

and

$$R^2 \Longrightarrow M,$$

where D and M are two random variables.

Spurious Regressions

- Macroeconomic time series are non-stationary and persistent
- Are all regressions in macroeconomics spurious then?!?!?!?!
- (...)

Co-integration

Integration

Definition ($I\left(d\right)$)

A series with no deterministic components which has a stationary, invertible, ARMA representation after differencing d times, is said to be integrated of order d, denoted $x_t \sim I\left(d\right)$.

- In this course, only the values $d = 0$ and $d = 1$ will be considered (e.g. random walk), but many results can be generalized to other cases including the fractional difference model
- Substantial differences between $x_t \sim I\left(0\right)$ and $x_t \sim I\left(1\right)$. (Random walk!)

Integration

Properties of $x_t \sim I(0)$ (with zero mean):

(i) $var(x_t)$ is finite

(ii) innovations have only temporary effect on x_t

(iii) autocorrelations, ρ_k, decrease steadily in magnitude for large enough k, so sum is finite

Integration

Properties of $x_t \sim I(1)$ ($x_0 = 0$):

(i) $var(x_t)$ goes to infinity as t goes to infinity

(ii) innovations have permanent effect on x_t

(iii) autocorrelations, $\rho_k \to 1$ for all k as $t \to \infty$

Co-integration

- If x_t and y_t are both $I(d)$, then it is "**generally"** true that the linear combination

$$z_t = x_t - ay_t,$$

will also be $I(d)$

- David Hendry said to Clive Granger: The difference of two $I(1)$ variables can be $I(0)$...

Co-integration

Engle, R. F. and C. W. J. Granger (1987): "Co-integration and Error Correction: Representation Estimation and Testing," *Econometrica* 55, 251-276.

Definition (Co-integration)

The components of the vector x_t are said to be co-integrated of order d, b, denoted $x_t \sim CI(d, b)$, if (i) all components of x_t are $I(d)$; (ii) there exists a vector α $(\neq 0)$ so that $z_t = \alpha' x_t \sim I(d-b), b > 0$. The vector α is called the co-integrating vector

Co-integration

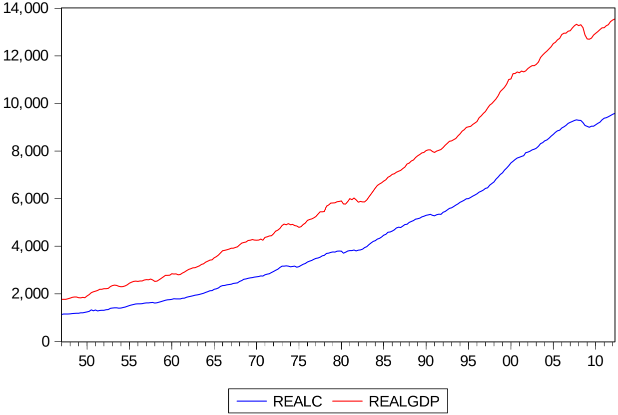
- Consider: $d = b = 1$. That is $x_t \sim I(1)$ and $z_t = \alpha' x_t \sim I(0)$
- **Simplest example**: Let $y_t \sim I(1)$, $x_t \sim I(1)$, and

  $$y_t = \theta x_t + z_t$$

  with $z_t \sim I(0)$
- Hence, z_t will rarely drift far from zero (if it has zero mean) and will often cross the zero line
- Equilibrium Relationship: $y_t - \theta x_t$; so that z_t represent the stationary deviation from the equilibrium
- Therefore, y_t and x_t will not move too far away from each other

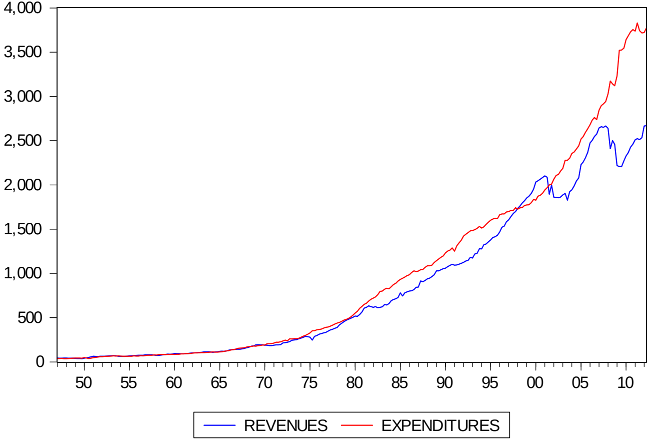
Co-integration

- Real Consumption and Real GDP
- U.S. Quarterly Data from the Federal Reserve Bank of St. Louis: 1947Q1-2012Q2



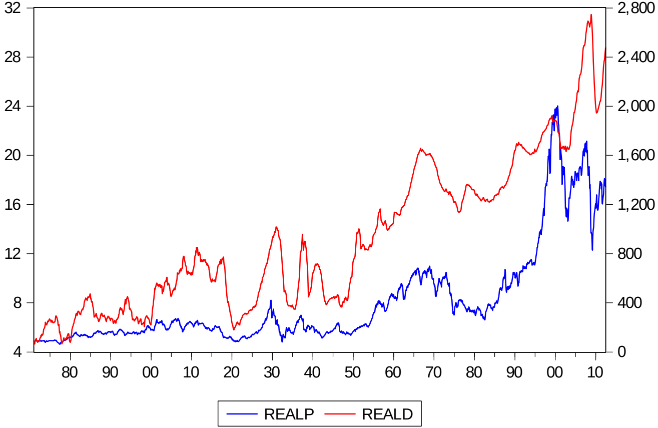
Co-integration

- Government Expenditures and Revenues
- U.S. Quarterly Data from the Federal Reserve Bank of St. Louis: 1947Q1-2012Q2



Co-integration

- Real Stock Prices and Real Dividends
- U.S. Monthly Data from Robert Shiller: 1871m1-2012m6



Co-integration

The common factor explanation

- Example: Let $W_t \sim I(1)$ and consider the following system

$$\begin{aligned} y_t &= aW_t + u_{yt} \\ x_t &= W_t + u_{xt} \end{aligned}$$

where u_{yt} and u_{xt} are both $I(0)$

- Then, $y_t \sim I(1)$ and $x_t \sim I(1)$
- But

$$y_t - ax_t = aW_t + u_{yt} - aW_t + au_{xt} = u_{yt} + au_{xt} \sim I(0)$$

- Cancellation of the common factor W_t!

Testing for Cointegration

Engle and Granger proposal

- Estimate the long run (static) relationship by OLS

$$y_t = f(t) + \beta x_t + z_t,$$

  and compute the OLS residuals, $\hat{z}_t$
- Then test for co-integration

$$\begin{cases} H_o : \hat{z}_t \sim I(1) \\ H_a : \hat{z}_t \sim I(0) \end{cases}$$

- **DF test on** $\hat{z}_t$: *Important!* Critical values are not the same as those derived by Dickey and Fuller ($\hat{z}_t$ incorporates the OLS estimates!)

- Critical values tabulated by **MacKinnon (1991)**

Testing for Cointegration

Estimation Properties

$$y_t = \theta x_t + z_t$$

In general,

- OLS estimates biased and inefficient (although super-consistent!)
- Nonstandard distributions with nuisance parameters
- There could be more than one cointegrating vector
- Endogeneity
- Static Relation

Single Equation <u>Dynamic</u> Models:

ARDL & ECM

Single Equation Dynamic Models

- Distributed Lags
- Autoregressive
- ARDL

Single Equation Dynamic Models

Distributed Lags

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + ... + \beta_r x_{t-r} + \varepsilon_t$$

with $E(\varepsilon_t | x_t, x_{t-1}, ...) = 0$.

(Multipliers + Graph)

Single Equation Dynamic Models

Autoregressive

$$y_t = \mu + \beta_0 x_t + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \varepsilon_t$$

with $E(\varepsilon_t | y_{t-1}, y_{t-2}, ... x_t) = 0$.

(Multipliers + Graph)

Single Equation Dynamic Models

ARDL

$$y_t = \mu + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + ... + \beta_r x_{t-r} + \varepsilon_t$$

with $E(\varepsilon_t | y_{t-1}, y_{t-2}, ...x_t, x_{t-1}, ...) = 0.$

(Multipliers + Graph)

Single Equation Dynamic Models

- Distributed Lags

$$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + ... + \beta_r x_{t-r} + \varepsilon_t$$

with $E(\varepsilon_t | x_t, x_{t-1}, ...) = 0$.

- Using the lag operator

$$y_t = \mu + (\beta_0 + \beta_1 L + ... + \beta_r L^r) x_t + \varepsilon_t$$

- Or more compactly

$$y_t = \mu + B_r(L) x_t + \varepsilon_t$$

Single Equation Dynamic Models

- Autoregressive

$$y_t = \mu + \beta_0 x_t + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \varepsilon_t$$

with $E(\varepsilon_t | y_{t-1}, y_{t-2}, ... x_t) = 0$.

- Using the lag operator

$$(1 - \gamma_1 L - ... - \gamma_p L^p) y_t = \mu + \beta_0 x_t + \varepsilon_t$$

- Or more compactly

$$C_p(L) y_t = \mu + \beta_0 x_t + \varepsilon_t$$

Single Equation Dynamic Models

- ARDL

  $$y_t = \mu + \gamma_1 y_{t-1} + ... + \gamma_p y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + ... + \beta_r x_{t-r} + \varepsilon_t$$

  with $E(\varepsilon_t | y_{t-1}, y_{t-2}, ...x_t, x_{t-1}, ...) = 0$.

- Using the lag operator

  $$(1 - \gamma_1 L - ... - \gamma_p L^p) y_t = \mu + (\beta_0 + \beta_1 L + ... + \beta_r L^r) x_t + \varepsilon_t$$

- Or more compactly

$$C_p(L) y_t = \mu + B_r(L) x_t + \varepsilon_t$$

Single Equation Dynamic Models

- IMPORTANT: ARDL models can be written in **Error Correction Form**
- Example: ARDL(1,1)

$$y_t = \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

- **ECM representation**

$$\Delta y_t = \mu + (\gamma_1 - 1)(y_{t-1} - \theta x_{t-1}) + \beta_0 \Delta x_t + \varepsilon_t$$

where $\theta = -(\beta_0 + \beta_1)/(\gamma_1 - 1)$

Single Equation Dynamic Models

- **Stability**: The ARDL model is stable when the roots of the autoregressive polynomial $C_p(L)$ are outside the unit circle

- If the model is stable then the following RD representation is well defined

$$y_t = \frac{\mu}{C_p(L)} + \frac{B_r(L)}{C_p(L)} x_t + \frac{1}{C_p(L)} \varepsilon_t$$

- $B_r(L)/C_p(L) = D_\infty(L)$ is convergent if the model is stable (finite impact of a change in x_t - back to equilibrium)

- Note that the new error term, say $u_t = \varepsilon_t / C_p(L)$ is now autocorrelated

- $\alpha = \mu / C_p(L) = \mu / C_p(1)$ is also finite

Single Equation Dynamic Models

- **DL representation of an ARDL model**

$$y_t = \frac{\mu}{C_p(L)} + \frac{B_r(L)}{C_p(L)} x_t + \frac{1}{C_p(L)} \varepsilon_t$$

- We can write this as

$$y_t = \alpha + D_\infty(L) x_t + u_t = \alpha + \sum_{j=0}^{\infty} \delta_j x_{t-j} + u_t$$

where $\alpha = \mu / C_p(1)$ and $u_t = \varepsilon_t / C_p(L)$

Single Equation Dynamic Models

Multipliers and Transmission of effects

- Impact (or contemporaneous) Multiplier:

$$m_0 = \frac{\partial y_t}{\partial x_t} = D_\infty(0) = \delta_0 = \frac{B_r(0)}{C_p(0)} = \beta_0$$

- j-th lag Multiplier:

$$m_j = \frac{\partial y_t}{\partial x_{t-j}} = \delta_j \neq \beta_j$$

- Total (or long run) Multiplier:

$$m_T = \sum_{j=0}^{\infty} m_j = D(1) = \sum_{j=0}^{\infty} \delta_j = \frac{B_r(1)}{C_p(1)}$$

Single Equation Dynamic Models

Measures of the Speed of Transmission

- Mean lag: tells us how concentrated or diluted the effect of x_t on y_t is. Earlier lags get higher weights
- Median lag: time when the dependent variable y_t has accumulated 50% of the total effect. If the effect is transmitted slowly, the Median Lag will be larger

Single Equation Dynamic Models

Measures of the Speed of Transmission

- Mean lag: weighted mean of all multipliers

$$MeanLag = \frac{\sum_{j=0}^{\infty} j\delta_j}{\sum_{j=0}^{\infty} \delta_j} = \frac{D'(1)}{D(1)} = \frac{B'(1)}{B(1)} - \frac{C'(1)}{C(1)}$$

where $D'(1) = \frac{dD(L)}{dL}|_{L=1}$. The mean lag tells us how concentrated or diluted the effect of x_t on y_t is. Earlier lags get higher weights

Single Equation Dynamic Models

Measures of the Speed of Transmission

- Median lag: time when the dependent variable y_t has accumulated 50% of the total effect

$$MedianLag = \min_q \left\{ \frac{\sum_{j=0}^{q} \delta_j}{\sum_{j=0}^{\infty} \delta_j} \geq 0.5 \right\}$$

If the effect is transmitted slowly, the Median Lag will be larger

Single Equation Dynamic Models

Correlation among regressors and error term

- Mean independence of the error term and the regressors is a fundamental assumption for OLS to deliver unbiased and consistent estimates

- This occurs when the error term is autocorrelated
- To see this, notice that

$$\begin{aligned} y_t &= \mu + \beta x_t + \alpha y_{t-1} + \varepsilon_t \\ y_{t-1} &= \mu + \beta x_{t-1} + \alpha y_{t-2} + \varepsilon_{t-1} \end{aligned}$$

- **Endogeneity!**

Single Equation Dynamic Models

Correlation among regressors and error term

- Example: AR(1) errors

$$\begin{aligned} y_t &= \mu + \beta x_t + \alpha y_{t-1} + \varepsilon_t \\ \varepsilon_t &= \phi \varepsilon_{t-1} + u_t \end{aligned}$$

where $u_t \sim i.i.d.(0, \sigma_u^2)$

- Example: MA(1) errors

$$\begin{aligned} y_t &= \mu + \beta x_t + \alpha y_{t-1} + \varepsilon_t \\ \varepsilon_t &= u_t + \theta u_{t-1} \end{aligned}$$

where $u_t \sim i.i.d.(0, \sigma_u^2)$

Single Equation Dynamic Models

Correlation among regressors and error term

- Endogeneity!
- What could we do?

Correlation among regressors and error term

- Endogeneity!
- What could we do?
- IV / 2SLS

Economic Hypotheses and Dynamic Models

- Adaptive Expectations
- Partial Adjustment

Single Equation Dynamic Models

Adaptive Expectations Hypothesis

- Consider

$$\begin{aligned} y_t &= \alpha + \beta x^e_{t+1|t} + \varepsilon_t \\ x^e_{t+1|t} &= \lambda x^e_{t|t-1} + (1-\lambda) x_t \end{aligned}$$

  where $x^e_{t+1|t}$ is the expectation at time t of x_{t+1}, $x^e_{t|t-1}$ is the previous expectation, x_t is the realization at time t, and $\lambda \in [0, 1]$

- When $\lambda = 0$: immediate correction. When $\lambda = 1$: expectations are static

Adaptive Expectations Hypothesis

- The expectations formation equation can be written as

$$x^e_{t+1|t} - x^e_{t|t-1} = (1-\lambda)(x_t - x^e_{t|t-1})$$

where $x^e_{t+1|t} - x^e_{t|t-1}$ is the change in the expectation;
$(1-\lambda)$ intensity of speed of adjustment of expectations;
$(x_t - x^e_{t|t-1})$ error in the last expectation formation

Adaptive Expectations Hypothesis

The expectations equation

$$x^e_{t+1|t} = \lambda x^e_{t|t-1} + (1-\lambda)x_t$$

can also be rewritten as

$$(1-\lambda L)x^e_{t+1|t} = (1-\lambda)x_t$$

Hence:

$$x^e_{t+1|t} = \frac{(1-\lambda)}{(1-\lambda L)}x_t = (1-\lambda)(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \ldots)$$

Adaptive Expectations Hypothesis

- The model is

$$\begin{aligned} y_t &= \alpha + \beta x^e_{t+1|t} + \varepsilon_t \\ x^e_{t+1|t} &= \lambda x^e_{t|t-1} + (1-\lambda) x_t \end{aligned}$$

where $x^e_{t+1|t}$ is the expectation at time t of x_{t+1}, $x^e_{t|t-1}$ is the previous expectation, x_t is the realization at time t, and $\lambda \in [0,1]$

- And we have seen that

$$x^e_{t+1|t} = \frac{(1-\lambda)}{(1-\lambda L)} x_t = (1-\lambda)(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + ...)$$

- Therefore, we can write the model in ADRL form

$$y_t = \alpha + \beta \frac{(1-\lambda)}{(1-\lambda L)} x_t + \varepsilon_t$$

Single Equation Dynamic Models

Adaptive Expectations Hypothesis

- Therefore, we can write the model in DL form

$$y_t = \alpha + \beta \frac{(1-\lambda)}{(1-\lambda L)} x_t + \varepsilon_t$$

- We can also write

$$(1-\lambda L) y_t = (1-\lambda L)\alpha + \beta(1-\lambda) x_t + (1-\lambda L)\varepsilon_t$$

or equivalently

$$y_t = \alpha^* + \beta^* x_t + \gamma^* y_{t-1} + \varepsilon_t^*$$

where $\alpha^* = (1-\lambda)\alpha$, $\beta^* = \beta(1-\lambda)$, $\gamma^* = \lambda$, and $\varepsilon_t^* = \varepsilon_t - \lambda\varepsilon_{t-1}$. Notice: $y_t \sim ARDL(1,0)$ with $\varepsilon_t^* \sim MA(1)$

Single Equation Dynamic Models

Partial Adjustment

- In Partial Adjustment models there is a target variable y_t^* that depends on another variable x_t; that is

$$y_t^* = \mu + \beta x_t + \varepsilon_t$$

- Partial Adjustment Hypothesis:

$$y_t - y_{t-1} = \gamma(y_t^* - y_{t-1})$$

  where $y_t - y_{t-1}$ is the observed change and $(y_t^* - y_{t-1})$ is the targeted change

- If $\gamma = 0$: there is no adjustment. If $\gamma = 1$: we get an instantaneous adjustment

Single Equation Dynamic Models

Partial Adjustment

- Partial Adjustment Hypothesis:
$$y_t - y_{t-1} = \gamma(y_t^* - y_{t-1})$$
where $y_t - y_{t-1}$ is the observed change and $(y_t^* - y_{t-1})$ is the targeted change

- Motivation: An agent is trying to minimize the following cost function:
$$c_t = a_1(y_t - y_{t-1})^2 + a_2(y_t^* - y_t)^2$$
where $(y_t - y_{t-1})$ is the adjustment cost and $(y_t^* - y_t)$ is the cost of not being optimal

- The solution to this optimization problem (based on the FOC $\partial c_t / \partial y_t = 0$) gives
$$y_t - y_{t-1} = \frac{a_2}{a_1 + a_2}(y_t^* - y_{t-1})$$

Single Equation Dynamic Models

Partial Adjustment

- Notice that the adjustment equation
$$y_t - y_{t-1} = \gamma(y_t^* - y_{t-1})$$
can be written as
$$y_t = \frac{\gamma}{(1-(1-\gamma)L)} y_t^*$$
- Partial Adjustment Hypothesis:where $y_t - y_{t-1}$ is the observed change and $(y_t^* - y_{t-1})$ is the targeted change
- Hence, back to the original model, that is
$$y_t^* = \mu + \beta x_t + \varepsilon_t$$
we see that we can then substitute y_t^* so that we get
$$\frac{(1-(1-\gamma)L)}{\gamma} y_t = \mu + \beta x_t + \varepsilon_t$$

Partial Adjustment

- Hence, back to the original model, that is

$$y_t^* = \mu + \beta x_t + \varepsilon_t$$

we see that we can then substitute y_t^* so that we get

$$\frac{(1-(1-\gamma)L)}{\gamma} y_t = \mu + \beta x_t + \varepsilon_t$$

- The model has an ARDL representation

$$y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 y_{t-1} + v_t$$

where $\alpha_0 = \gamma\mu$, $\alpha_1 = \gamma\beta$, $\alpha_2 = (1-\gamma)$, and $v_t = \gamma\varepsilon_t$

Co-integration and Dynamic Models

- We have seen that ARDL models can be written in ECM form
- There is a very important result that connects the concept of co-integration with the ECM representation
- This result is known as the **Granger Representation Theorem**
- The theorem says (roughly) that a vector of $I(1)$ time series is co-integrated if and only if an error correction representation exists.
- This theorem is of great value both in theory and in practice.

Co-integration and Dynamic Models

By the **Granger Representation Theorem**, two time series y_t and x_t are co-integrated if and only if the following **ECM representation** exist

$$\begin{aligned}
\Delta y_t &= -\gamma_1(y_{t-1} - \beta x_{t-1}) + \Phi_1(L)\Delta y_{t-1} + \Omega_1(L)\Delta x_{t-1} + \varepsilon_{1t} \\
\Delta x_t &= -\gamma_2(y_{t-1} - \beta x_{t-1}) + \Phi_2(L)\Delta y_{t-1} + \Omega_2(L)\Delta x_{t-1} + \varepsilon_{2t}
\end{aligned}$$

Co-integration and Dynamic Models

ECM

$$
\begin{aligned}
\Delta y_t &= -\gamma_1(y_{t-1} - \beta x_{t-1}) + \Phi_1(L)\Delta y_{t-1} + \Omega_1(L)\Delta x_{t-1} + \varepsilon_{1t} \\
\Delta x_t &= -\gamma_2(y_{t-1} - \beta x_{t-1}) + \Phi_2(L)\Delta y_{t-1} + \Omega_2(L)\Delta x_{t-1} + \varepsilon_{2t}
\end{aligned}
$$

- $-\gamma_1(y_{t-1} - \beta x_{t-1})$: error correction term
- γ_1: speed of adjustment
- $y_{t-1} - \beta x_{t-1}$ long-run equilibrium relationship

Co-integration and Dynamic Models

ECM

$$
\begin{aligned}
\Delta y_t &= -\gamma_1(y_{t-1} - \beta x_{t-1}) + \Phi_1(L)\Delta y_{t-1} + \Omega_1(L)\Delta x_{t-1} + \varepsilon_{1t} \\
\Delta x_t &= -\gamma_2(y_{t-1} - \beta x_{t-1}) + \Phi_2(L)\Delta y_{t-1} + \Omega_2(L)\Delta x_{t-1} + \varepsilon_{2t}
\end{aligned}
$$

- $\Phi_1(L)\Delta y_{t-1} + \Omega_1(L)\Delta x_{t-1}$: short run dynamics
- the model could include deterministic terms

Co-integration and Dynamic Models

Co-integration is a facinating theory!

We could continue talking about it for another term,

or two, or three... but...

The End

<u>End of the course</u>

I hope you have enjoyed it

and

learnt some econometrics along the way...

$Merry\ \widehat{Christmas}$!