



UNIVERSITEIT•STELLENBOSCH•UNIVERSITY  
jou kennisvenoot • your knowledge partner

# Master's Applied Time-Series Econometrics

## Session 4: Modelling Long-Run Relationships

**Lecturer:**

Nico Katzke

[nicokatzke@sun.ac.za](mailto:nicokatzke@sun.ac.za)



Department of Economics  
(copyright)

DEPARTMENT OF  
ECONOMICS



# What we will be discussing today

---



- Today we are looking at applying Co-Integration tests and VECMs.
- These are key tools in analysing long term dynamic relationships between different time-series.
- We will be covering the theory today and thereafter we will look at these techniques applied on local asset markets.



# VAR recap



- Remember from last time, the VAR(I) model of  $x_t$  and  $y_t$  can be written as:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{x_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Or:

$$x_t = \alpha_{11}x_{t-1} + \alpha_{12}y_{t-1} + \varepsilon_{x_t}$$

$$y_t = \alpha_{21}x_{t-1} + \alpha_{22}y_{t-1} + \varepsilon_{y_t}$$

We then used OLS to estimate the model, by assuming that  $\varepsilon_{x_t}$

&  $\varepsilon_{y_t}$  are uncorrelated with  $x_t$  &  $y_t$



# VAR recap



- From the model we can get one period ahead forecasts as:

$$E_t(x_{t+1}) = \alpha_{11}x_t + \alpha_{12}y_t$$

$$E_t(y_{t+1}) = \alpha_{21}x_t + \alpha_{22}y_t$$

- We then iteratively forecast into the future based on estimates for inputs. E.g. forecasting to  $t + 2$  from period  $t$ :

$$E_t(x_{t+2}) = \alpha_{11}E_t(x_{t+1}) + \alpha_{12}E_t(y_{t+1})$$

$$E_t(y_{t+2}) = \alpha_{21}E_t(x_{t+1}) + \alpha_{22}E_t(y_{t+1})$$

Which gets inaccurate as the forecast horizon increases – as the inputs are not perfect (are themselves estimates) and the parameters are not perfect (they're assumed constant over time and are estimated based on past data).



# Forecast error variance & Granger Causality



- We then calculate the one-period forecast errors as:

$$x_{t+1} - E_t(x_{t+1}) = \varepsilon_t$$

$$x_{t+2} - E_t(x_{t+2}) = \varepsilon_{t+1}$$

And from these we can calculate the one-step ahead forecast error variance.

Refer back to the first semester for a discussion on fitting and using the forecast error variances.

Another important definition covered was that of **Granger Causality**, where  $y_t$  is said to Granger cause  $x_t$  if  $y$  helps to predict  $x$ . This might not at all imply statistical causality, but is referring to the ability of lags of  $y$  to help in predicting contemporary values of  $x$ . Thus from the earlier notation, whether  $\alpha_{21}$  is significant or not (we typically test more than one lag though...)



# Co-Integration testing

---



- In addition to using cointegration techniques as a means of doing forecasting, studying mean-reversion or explaining long run behaviour,
- Establishing cointegration is also a starting point for most long-term relationship studies in much of macroeconomic and financial research:

***If the returns are co-integrated, it warrants a further study into the dynamics of the relationship between the series...If two series are not cointegrated, it may show that these series are not counterinfluencing each other over time.***



# Co-Integration testing

---



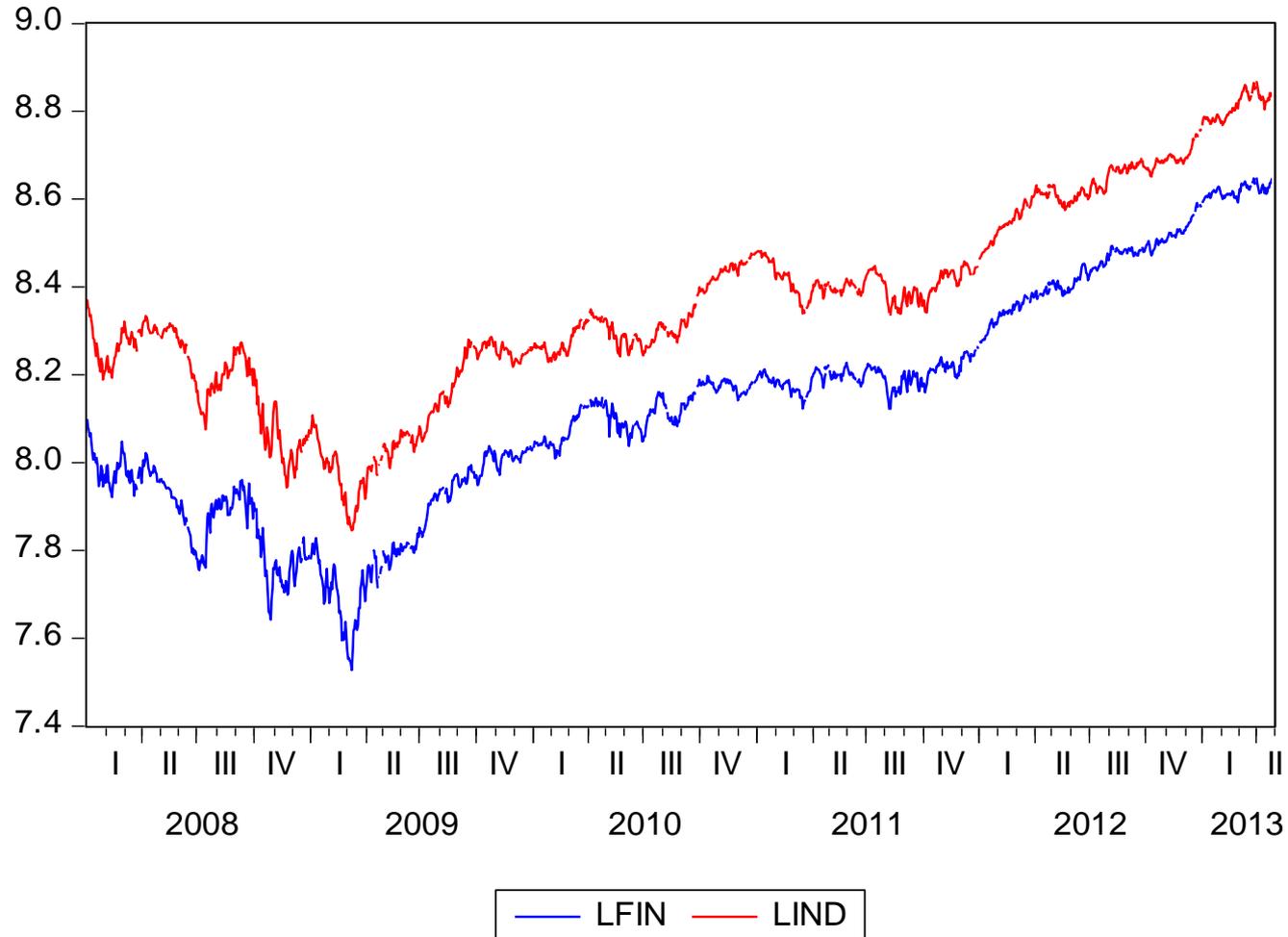
- Basically, Co-integration studies whether a set of series display co-movements over time and have the same order of integration, so that they share a common stochastic trend – so that when we control for the long run path the series follow, we might have a time-invariant series...
- **This implies that there exists a long run relationship between the set of variables that is stationary.**
- To illustrate this, consider the **Industrials** and **Financials** industries for South Africa since 2008



# Co-integration



- \*Note: we use the logarithm of closing prices for both

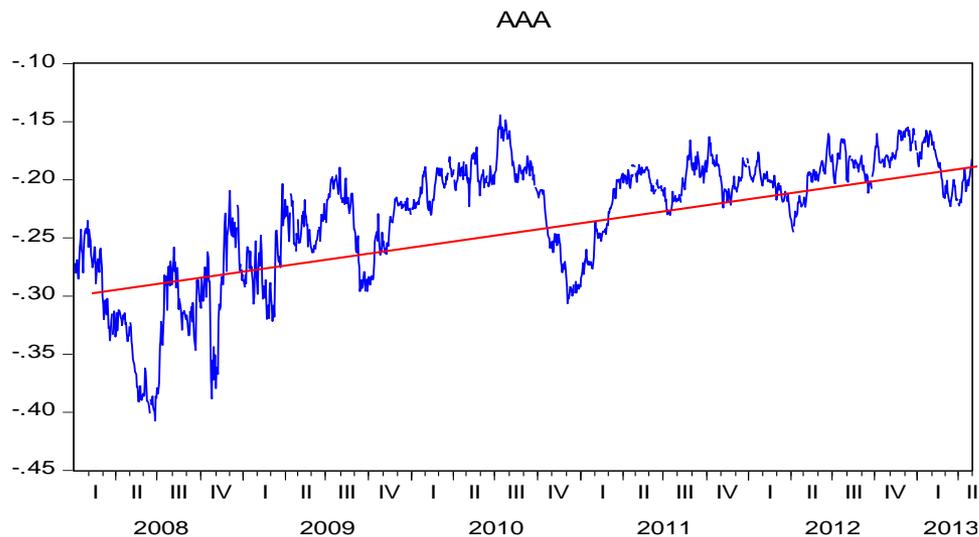




# Co-integration



- From the graph it is clear that both  $d\text{fin}$  and  $d\text{lind}$  have unit roots (or, more formally, it strongly rejects the ADF and Peron test for stationarity).
- But when we take the difference, it seems that the long run relationship **may be stationary** (although in this case there may be a constant trend component (red-line) too...)





# Co-integration



- Why would we do this? Remember that using  $I(1)$  variables in a regression framework leads to **spurious results** so we cannot do that. Also, using first differences removes a lot of NB information.
- So for both  $d\text{fin}$  and  $d\text{lind}$  being  $I(1)$ , it is clear that the combination:  
$$[ I\text{Fin}_t - \beta d\text{lind}_t - k.t ] \sim I(0)$$
- (Where  $t$  is the time-trend component, and  $\beta$  the LR relationship parameter)

**Thus we can define co-integration as:** the combination of  $I(1)$  variables that have a linear combination that is  $I(0)$   $\rightarrow$  which implies the series evolve according to a long-run path, with  $I(0)$  deviations from it in the short term, but ultimately it reverts back to the LR path.



# Usefulness...



- Now obviously this is very useful, as understanding this long term relationship and the short run dynamics, we can better model future asset market behaviour!

\*Note that we need to **normalize** one variable so that meaningful (relative) analysis can be made.

Hence, suppose the cointegrating relationship is:

$$\beta_1 x_t - \beta_2 y_t \sim I(0)$$

Then we would use:  $x_t - \frac{\beta_1}{\beta_2} y_t \sim I(0)$



# Common Factor:

$X_t$  &  $Y_t$  need not be ***Directly*** linked

---



- Often times, the source of LR commonality might be due to a **common external factor** (consider e.g. the Industrials and Financials series – the LR path might be driven by common market factors).
- Thus Suppose  $X_t$  &  $Y_t$  are  $I(1)$  and are driven by:

$$Y_t = \mu_0 + \lambda \cdot Z_t + \varepsilon_t$$

$$X_t = \mu_1 + Z_t + v_t$$

With:  $Z_t \sim I(1)$  &  $\varepsilon_t$  &  $v_t \sim I(0)$

- Then:

$$Y_t - \lambda \cdot X_t = \{\mu_0 - \lambda(\mu_1)\} + e_t - \lambda v_t$$

With  $Y_t - \lambda \cdot X_t \sim I(0)$  **by definition (why?)**, despite  $Y_t \rightarrow \neq f(X_t)$



# Natural Cointegration:



- We could also have the two equations directly linked:

$$Y_t = \beta_0 + \beta_1 Z_t + \epsilon_t \quad (1)$$

With:  $\epsilon_t \sim I(0)$  &  $Z_t \sim I(1)$  given as:

$$Z_t = Z_{t-1} + \phi + v_t \quad (2), \quad v_t \sim I(0)$$

Then differencing (1):

$$\Delta Y_t = \beta_1 \Delta Z_t + \Delta \epsilon_t$$

$$\Delta Y_t = \beta_1 (\phi + v_t) + \Delta \epsilon_t$$

Thus:

$$Y_t = Y_{t-1} + \gamma + \mu_t, \text{ with:}$$

$$\gamma = \beta_1 (\phi), \quad \mu_t = \Delta \epsilon_t + \beta_1 v_t$$

**So that it follows:**

$$Y_t \sim I(1), \text{ but } (Y_t - \beta_1 Z_t) = \sum \{I(0)\text{series}\} = I(0)$$



# Fitting it in Eviews



- We will use the Johansen Linear Cointegration test in Eviews to establish cointegration:
- It works as follows. Suppose we have:

$k$  variables  $y_1 \dots y_k$  in vector form,  $Y_t$ , that are all  $I(1)$ .

- Johansen test then estimates a Multivariate ADF test using the following VAR(p) as input (assume a VAR(2) for the system):

$$Y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \gamma(\phi_t) + \varepsilon_t$$

Which can be rewritten as: **(if we  $\pm Y_{\{t-1\}}$  &  $\beta_2 Y_{\{t-1\}}$  both sides)**

$$\Delta Y_t = (\beta_1 + \beta_2 - I_k) y_{t-1} - \beta_2 y_{t-1} + \beta_2 y_{t-2} + \gamma(\phi_t) + \varepsilon_t,$$

...which is:  $\beta_2 \cdot (\Delta Y_{t-1})$



# Fitting it in Eviews



With  $Y_t$  the  $k \times 1$  vector of  $I(1)$  variables and  $\phi_t$  the vector of deterministic variables (trend components, etc).

This is then written as:

$$\Delta Y_t = \pi \cdot y_{t-1} + \delta_1 \cdot \Delta y_{t-1} + \gamma(\phi_t) + \varepsilon_t,$$

With:

$$\pi = (\beta_1 + \beta_2 - I_k) \quad \& \quad \delta_1 = -\beta_2$$

So that in the more general form (p-lags) we have:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Delta y_{t-i} + \gamma(\phi_t) + \varepsilon_t$$



# Simplifying somewhat:



- Ignoring deterministic components (assume no trends) and no lags, then rewrite the previous slide as:

$$Y_t = \beta_1 y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \beta \cdot y_{t-1} - y_{t-1} + \varepsilon_t \quad \dots (+ \setminus - y_{t-1} \text{ both sides})$$

- Thus: 
$$\Delta y_t = (\beta - 1) \cdot y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \pi \cdot y_{t-1} + \varepsilon_t$$

- The  $\Pi$ -matrix thus takes the form as:

$$\Pi = \sum_{i=1}^p (\beta_i - I)$$



# Pi in the Sky?!



- So what is so great about this  $\pi$  matrix?

$$\text{From: } \Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Delta y_{t-i} + \gamma(\phi_t) + \varepsilon_t$$

We see that  $\pi$  determines how the level process of  $y_t$  enters the **dynamic process**  $\Delta y_t$ .

When  $\Pi = 0$ , the dynamic system does not depend on the levels of any process...

Thus:  $\Pi = \sum_{i=1}^p (\beta_i - I)$  is the matrix of Long Run responses ::::: i.e.

how incremental changes are affected by the LR level...



# Rank zero



$$\Pi = \sum_{i=1}^p \beta_i - I$$

Remember from the first session on matrixes:

*Rank( $\pi$ ) = Number of linearly independent rows & columns.*

Thus if the *Rank( $\pi$ ) = 0*, then all  $\beta$ 's  $\rightarrow$  and  $\beta$  becomes the Identity matrix, and we have:

$$\Delta y_t = (I - I) \cdot y_{t-1} + \varepsilon_t$$

Thus:

$$\Delta y_t = \varepsilon_t$$

Which is a first order VAR process... and **all the  $y_t$  series have unit roots, with no linear combination yielding a stationary process.**



# Rank Full



- Conversely, assume  $Rank(\pi) = k$ .
- Now the long run solution looks as follows:

$$\pi_{11}Y_{1t} + \pi_{12}Y_{2t} + \dots + \pi_{1k}Y_{kt} = 0$$

...

...

$$\pi_{k1}Y_{1t} + \pi_{k2}Y_{2t} + \dots + \pi_{kk}Y_{kt} = 0$$

So that:

- Our long run solution now has  $k$  – independent stationary solutions... THUS all variables are  $I(0)$ !!



# Fitting it in Eviews



$$\Pi = \sum_{i=1}^p \beta_i - I$$

Granger's theorem centres around finding the **rank of  $\pi$**  that yield stationarity of the  $y_t$  process...

Hence a Multi-Variate ADF test effectively!

Thus if the coefficient matrix  $\Pi$  has a **reduced rank** ( $r < k$ ) **→ then** there exists ( $k \times r$ ) matrices of rank  $r$  **such that:**  $(\beta' Y_t) \sim I(0)$ .

Thus  $r$  is the **cointegration rank**, implying that:

**Each column of  $\beta$  in  $\beta' Y_t$  is the cointegration vector (and they're linearly independent columns).**



# Understanding what's going on



- The key to understanding what is going on in cointegration analysis is to understand the matrix that governs the LR cointegration relationship:  $\Pi$  can always be **decomposed into two matrices**:

$$\Pi = \alpha\beta'$$

With  $\alpha$  &  $\beta$  both  $k \times r$  matrices, with  $r$  being the number of linear cointegrating relationships (i.e.  $\text{rank} = r$ ).

- Thus the **number of Linear LR cointegrating relationships** for the  $k$  variables in  $Y_t = \text{Rank}(\Pi) = r$ .
- ***Which means there are  $r$  linearly independent combinations of the variables in  $Y_t$  that yield a stationary  $I(0)$  process...***
- As will be seen:
  - $\alpha$  = speed of adjustment parameters.
  - $\beta$  = cointegrating vectors.



# Cointegration testing in Eviews:



- Select the series of interest.
- Make sure that they are all  $I(1)$  (sometimes in the literature variables that show **very high persistence yet do not reject the ADF** are included too... (why?))
- Highlight the group – Quick / Johansen Cointegration test.
- The following appears:

Johansen Cointegration Test

Cointegration Test Specification

Deterministic trend assumption of test—

Assume no deterministic trend in data:

1) No intercept or trend in CE or test VAR

2) Intercept (no trend) in CE - no intercept in VAR

Allow for linear deterministic trend in data:

3) Intercept (no trend) in CE and test VAR

4) Intercept and trend in CE - no intercept in VAR

Allow for quadratic deterministic trend in data:

5) Intercept and trend in CE - intercept in VAR

Summary:

6) Summarize all 5 sets of assumptions

Exog variables\*

Lag intervals

1 4

Lag spec for differenced endogenous

Critical Values

MHM

Size 0.05

Osterwald-Lenum

\* Critical values may not be valid with exogenous variables; do not include C or Trend.

OK Cancel



# Cointegration choices:



From Johansen (1995:80):

1. The data in  $y_t$  have **no** deterministic trends and the cointegrating equations do not have any intercepts.:

$$H_0: \Pi Y_{t-1} + \gamma(\phi_t) = \alpha\beta'Y_{t-1} \quad (\text{with } \alpha\beta' \text{ s reduce rank tested})$$

2. The data in  $y_t$  have **no** deterministic trends while the cointegrating equations **do have** intercepts (and no trends).
3. The data in  $y_t$  **have** linear trends, while the cointegrating equations have **only** intercepts (and no trends).
4. The data in  $y_t$  as well as the cointegrating equations have linear trends.
5. The data in  $y_t$  have **quadratic** trends, while the cointegrating equations have linear trends.



# A note on the deterministic components



- We've focussed now on  $\pi$ , but we see we have to make a choice regarding  $\gamma(\phi_t)$  too...
- Our choice on the deterministic components govern the system's behaviour i.t.o. level drift & linear (or even quadratic trends in cointegrating vectors.
- Certain variables have no possibility of inherent drifts (e.g. interest rates), while in most cases quadratic trends do not make much sense – i.e. we should force the **time-trend** to lie within the cointegration space... **ELSE:**

$$\Delta Y_t = (\text{Coint Space}) + (t) \rightarrow \text{implying } Y_t \sim f(t^2) \dots?!!$$



# Cointegration choices:

---



- After selecting which description best fits the data (checking is NB), we need to decide on the amount of lags ( $p$ ) to include (default is  $k$  – amount of variables in the system).
- Also NB – of course we should use the **Levels** of the series as inputs (using Dlogs would defy the exercise right?! – remember we must **only** include  $I(1)$  series into the cointegration test).
- We then get the following output (suppose we chose default):



# Inlcuding 5 economic sectors in SA:



## Johansen Cointegration Test

Date: 07/30/13 Time: 16:07  
 Sample (adjusted): 1/09/2002 4/30/2013  
 Included observations: 2828 after adjustments  
 Trend assumption: Linear deterministic trend  
 Series: FINANCIALS INDUSTRIALS HEALTH\_CARE TELECOMMUNICATIONS BASIC\_MATERIALS  
 Lags interval (in first differences): 1 to 4

### Unrestricted Cointegration Rank Test (Trace)

| Hypothesized No. of CE(s) | Eigenvalue | Trace Statistic | 0.05 Critical Value | Prob.** |
|---------------------------|------------|-----------------|---------------------|---------|
| None *                    | 0.017415   | 111.6782        | 69.81889            | 0.0000  |
| At most 1 *               | 0.014077   | 61.99433        | 47.85613            | 0.0014  |
| At most 2                 | 0.004615   | 21.90234        | 29.79707            | 0.3039  |
| At most 3                 | 0.001844   | 8.821098        | 15.49471            | 0.3822  |
| At most 4                 | 0.001273   | 3.602815        | 3.841466            | 0.0577  |

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level →  
 \* denotes rejection of the hypothesis at the 0.05 level  
 \*\*MacKinnon-Haug-Michelis (1999) p-values

**Result**

### Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

| Hypothesized No. of CE(s) | Eigenvalue | Max-Eigen Statistic | 0.05 Critical Value | Prob.** |
|---------------------------|------------|---------------------|---------------------|---------|
| None *                    | 0.017415   | 49.68385            | 33.87687            | 0.0003  |
| At most 1 *               | 0.014077   | 40.09199            | 27.58434            | 0.0008  |
| At most 2                 | 0.004615   | 13.08124            | 21.13162            | 0.4448  |
| At most 3                 | 0.001844   | 5.218283            | 14.26460            | 0.7141  |
| At most 4                 | 0.001273   | 3.602815            | 3.841466            | 0.0577  |

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level  
 \* denotes rejection of the hypothesis at the 0.05 level  
 \*\*MacKinnon-Haug-Michelis (1999) p-values

Number of cointegrating relations under the null hypothesis that is being tested.

Eigenvalues of  $\Pi$

Either Trace or Max-Eigenvalue technique used for statistic.

Prob shows the p-value for the  $H_0: \text{Rank}(\Pi) \leq r$

If  $p < 0.05$  it implies that the cointegrating relationship is significant and the rank is at least as large as  $r$  (there is **at least**  $r$  linear combinations of variables that are stationary).



# Interpreting output



- **The output shows two statistics: Trace and Maximum Eigenvalue.**
- **Trace** → Tests  $H_0$ :  $r$  – cointegrating relationships exists for the  $k$  –variables, against the alternative of  $k$  (for  $r = 0, 1, \dots k - 1$ ). This is done sequentially starting at  $r = 0$ . Note that the alternative  $k$  (thus where  $\text{Rank}(\Pi) = k = \text{full rank}$  corresponds to the case where none of the series have a unit root and all are effectively  $I(0)$  → implying no need for cointegration testing, and a stationary VAR can be specified by the series in levels.
- **Max Eigenvalues** → it's a similar test, but this time reports the maximum eigenvalue statistic testing  $H_0$ :  $r$  – cointegrating relationships against alternative of  $r + 1$  – cointegrating relationships.



# Cointegrating relationships



- Thereafter all the cointegrating equations' normalized  $\alpha$  &  $\beta$  values are given...
- E.g. for the economic sectors example in SA:

|   |                |             |           |           |
|---|----------------|-------------|-----------|-----------|
| 1 Cointegrating Equation(s):  | Log likelihood | -41849.56   |           |           |
| Normalized cointegrating coefficients (standard error in parentheses) |                |             |           |           |
| FINANCIALS  | HEALTH_CARE    | INDUSTRIALS | RESOURCES | MINING    |
| 1.000000  | -0.105298      | -0.752309   | 4.000396  | -3.911401 |
|   | (0.05661)      | (0.10818)   | (0.84263) | (0.86798) |
| Adjustment coefficients (standard error in parentheses)               |                |             |           |           |
| D(FINANCIALS)   | 0.002810       |             |           |           |
|   | (0.00587)      |             |           |           |
| D(HEALTH_C...   | 0.018782       |             |           |           |
|   | (0.00667)      |             |           |           |
| D(INDUSTRIA...  | 0.015932       |             |           |           |
|   | (0.00654)      |             |           |           |
| D(RESOURC...  | -0.009482      |             |           |           |
|   | (0.00681)      |             |           |           |
| D(MINING)   | -0.008231      |             |           |           |
|   | (0.00673)      |             |           |           |

Cointegrating Relationship 1:

Normalized  $\beta$

$\alpha$



# Reporting:



- When reporting the cointegration test in your analysis, it is often best to provide the summary of the statistics by selecting option 6 in the output, giving you:

| Selected (0.05 level*) Number of Cointegrating Relations by Model |                          |                       |                       |                    |                    |
|---|--------------------------|-----------------------|-----------------------|--------------------|--------------------|
| Data Trend:   | None                     | None                  | Linear                | Linear             | Quadratic          |
| Test Type   | No Intercept<br>No Trend | Intercept<br>No Trend | Intercept<br>No Trend | Intercept<br>Trend | Intercept<br>Trend |
| Trace   | 3                        | 3                     | 3                     | 3                  | 2                  |
| Max-Eig   | 3                        | 3                     | 3                     | 3                  | 1                  |

\*Critical values based on MacKinnon-Haug-Michelis (1999)

- **Interpretation:**

*The output suggests that there exists **at least** one linear, non-trivial (i.e. where  $\beta \neq 0$ ), cointegrating relationship between the series included in the study, such that we can conclude that there exists a long term combination of these series that exhibit stationarity.*



UNIVERSITEIT•STELLENBOSCH•UNIVERSITY  
jou kennisvenoot • your knowledge partner

# VECMs



Department of Economics

DEPARTMENT OF  
ECONOMICS



- Having established a non-trivial linear combination of the  $I(1)$  – financial variables that produce a stationary residual series using the cointegrating relationships, **it warrants us to study the long-run relationship between them in depth.**
- One way we can do this is by including the Error Correction Models (ECM) component, or the variant VECM (Vector ECM).
- These models are restricted forms of VARs, designed **specifically** for use on series that are  $I(1)$  **and cointegrated.**



# ECM



- To illustrate the Error Correction Model process simply in a bivariate case, let's consider the Engle Granger approach to fitting an ECM.
- Suppose we have the Spot (S) and the Future (F) rate for corn.
- To fit an ECM, **we first have to establish** whether the series have unit roots (also, we would use the Log forms). Thus:  $S_t, F_t \sim I(1)$
- We then run a regression:

$$S_t = c + \beta \cdot F_t + \epsilon_t$$

And then we conduct an ADF test on  $\epsilon_t \rightarrow$  if  $\epsilon_t \sim I(0)$  it would imply that the series are **cointegrated** (a linear combination yields a stationary residual series...).

That is **step one**.



- Based on our estimation and thus identification of  $\beta$  in **step one**, we have established a cointegrating relationship **if** the residuals are indeed stationary.
- This implies that fitting an ECM would allow us to study the LR and SR dynamics between  $S_t$  &  $F_t$ :

$$\Delta(S_t) = c + \phi \cdot \Delta(F_t) + \gamma \cdot EC_{t-1} + \varepsilon_t$$

With:

$EC_t = (S_t - \beta \cdot F_t) \rightarrow$  Error correction term

$\phi \rightarrow$  describes the SR relationship between  $S_t$  &  $F_t$

$\gamma \rightarrow$  Describes the speed of adjustment process following deviations from the LR equilibrium, i.e. where  $S_t = \beta \cdot F_t$ .

$\varepsilon_t = WN$  (Check this)



# MAL



- From the ECM we can then compute the MAL (Mean Adjusted Lag) statistic (Doornik & Hendry, 1994), which describes how long it takes for the adjustment lag to return to equilibrium following a deviation:  $S_t \neq \beta \cdot F_t$  measured as:

$$MAL = (1 - \phi) / \gamma$$

So that if  $MAL$  is high, then the speed of adjustment is slow and *vice versa*.



# VECM



- The Engle-Granger VECM approach is great for studying the interdependent dynamics between two variables, with the general VECM form being (ignoring all deterministic terms such as trends):

$$\Delta(X_t) = c_1 + \phi_1 \cdot \Delta(Y_t) + \gamma_1 \cdot EC_{t-1} + \varepsilon_{1,t}$$

$$\Delta(Y_t) = c_2 + \phi_2 \cdot \Delta(X_t) + \gamma_2 \cdot EC_{t-1} + \varepsilon_{2,t}$$

$\phi$  – parameter of short run relation between:  $\Delta x$  &  $\Delta y$

$\gamma$  → speed of adjustment of deviation from long run trend.

$\beta$  → LR relation between  $x$  &  $y$  in cointegrating relationship  $x_t = \beta y_t$

- But when we have more than one cointegrating relationship, we use the Johansen framework for determining the cointegration relationships.
- The VECM then is based essentially on the same principles as the Engle-Granger two step ECM approach... yet can handle  $r > 1$ .



# VECM



- Granger demonstrated that **if** a system is cointegrated then there exists a VECM, and vice versa.
- Basically what the model does is it explains short run deviations from the long term trend / relationship.
- A simple form of VECM can be given by:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

Now, knowing that  $x$  &  $y$  are **cointegrated** (necessary condition), with :  
**cointegrating relationship:**  $(x_t = \beta y_t)$ , there then exists a matrix  $\Pi = \alpha\beta'$  such that:  $(x_t - \beta y_t) \sim I(0)$ .

We can rewrite the VECM as:  $\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} [1 - \beta] \cdot \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$



- This VECM above can thus be rewritten as:

$$\Delta x_t = \alpha_1(x_{t-1} - \beta y_{t-1}) + \varepsilon_{1,t}$$

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

Notice from the VECM that there is a deviation from the long-run trend  $(x_{t-1} - \beta y_{t-1})$  in  $(t - 1)$ . As in ECM this term we call the error correction term.

Now you can also see that  $\alpha$  represents the **speed** of adjustment to which deviations from the long run revert back to equilibrium.

*(Of course not all  $\alpha$ 's can be zero – else the system would not be stable. We can, however, impose restrictions on  $\alpha$ ).*

$\beta$  here is the **cointegrating vector** that represents the LR relation between  $x$  &  $y$  that is stationary.



# VECM



- Why the “error correction” in the name?
- Because we assume that, in the bivariate case,  $x_t = \beta y_t$  is a long run stationary relationship that holds over time.
- When there are deviations from this equilibrium:  $(x_t - \beta y_t)$ , the model adjusts at a speed of  $\alpha$  to bring the relationship back to the long run equilibrium – driving the error correction of  $x_t$ , or  $\Delta x_t$
- Note this implies we predict the change in  $x_t$  ( $\Delta x_t$ ) needed to bring us back to LR equilibrium.



# VECM – when to use?



- VECM's are rearranged and restricted forms of a VAR, appropriate for use with cointegrated time series.
  - *To check this, rewrite VAR(1):  $y_t = \phi_1 y_{t-1} + \varepsilon_t$  as the VECM:*
$$\Delta y_t = \alpha \beta' y_{t-1} + \varepsilon_t, \quad \text{with} \quad \alpha \beta' = \phi_1 - I_p$$
- Remember thus that the **first step** is ensuring all the variables are indeed  $I(1)$  and also cointegrated to **at least** the first order, and are less than full rank. Thus:  $r \in [1, k - 1]$ .
- Thus the sequence is:
  - Use ADF / Perron to establish the series are  $I(1)$ .
  - Use Johansen to show the series are cointegrated and determine rank of  $\pi$ .
  - Find  $\alpha$  &  $\beta$  in the VECM form of  $\Delta y_t = \alpha \beta' y_{t-1} + \varepsilon_t$  using OLS.



# VECM



- Suppose that there are  $g$  series and the cointegration rank is  $r$ .
- Then if  $r = 1$  and  $g = 4$ , the cointegrating matrix  $\pi$  looks like:

- $\pi = \alpha\beta' = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{bmatrix} \cdot [\beta_{11}\beta_{12}\beta_{13}\beta_{14}]$   LR part

- And for  $r = 2$ :  SR part

- $\pi = \alpha\beta' = \begin{bmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \\ \alpha_{13} & \alpha_{23} \\ \alpha_{14} & \alpha_{24} \end{bmatrix} \cdot \begin{bmatrix} \beta_{11}\beta_{12}\beta_{13}\beta_{14} \\ \beta_{21}\beta_{22}\beta_{23}\beta_{24} \end{bmatrix}$



UNIVERSITEIT•STELLENBOSCH•UNIVERSITY  
jou kennisvennoot • your knowledge partner

# Could or Should we use VECMs?



Department of Economics

DEPARTMENT OF  
ECONOMICS



## Could we or should we use VECM?



- If the series are **cointegrated &  $I(1)$** , that would imply we **should** fit a VECM and not a VAR to explain the variables' dynamics.
  - This follows as taking the First Differences of the time series and then fitting a VAR (thus a VAR of the Dlogged series), the resulting model would be **mis-specified** because of the **omitted error correction term**  $(x_{t-1} - \beta y_{t-1})$  for the series  $\Delta x_t$  (thereby ignoring the short run dynamic relationship between  $x$  &  $y$  following a deviation from the LR paths), while conversely running a VAR in **level forms** for the non-stationary  $I(1)$  series would yield spurious results.
- **THUS** in the presence of cointegration – **a VECM is required.**



# VECM: Deterministic terms

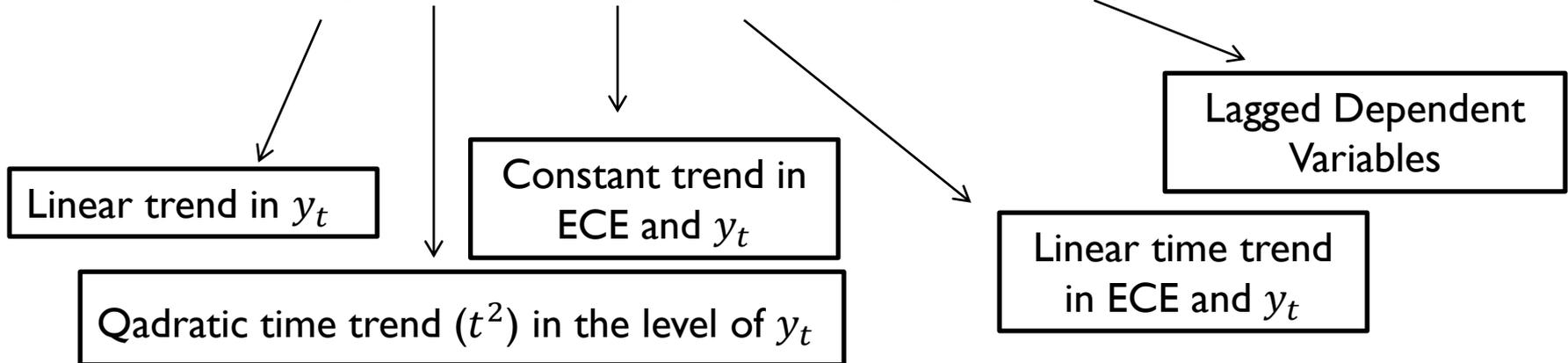


- Let's consider the inclusion of constant terms and time trends in a VECM of the series contained in the vector

$$y_t = (y_{1,t}, \dots, y_{N,t})$$

- Such deterministic terms can be included in the cointegrating equation (CE) or in the error correction equation (ECE).
- The general model can be written as:

$$\Delta y_t = \varphi_0 + \varphi_1 t + \alpha(\beta_0 + \beta_1 t + \beta' \mu_{t-1}) + \gamma \Delta y_{t-1} + \dots + \varepsilon_t$$





# VECM: Deterministic specifications



- Deciding on the inclusion of such terms, one needs to consider the relationship between the series of interest (there are 5 types).
- 1)  $\varphi_0 = \varphi_1 = 0$  &  $\beta_0 = \beta_1 = 0$ , implies the **series** have no trends, but may have non-zero means (but are removed by the difference), while the ECE has no constants.
- 2)  $\varphi_0 = \varphi_1 = 0$  &  $\beta_0 \neq 0, \beta_1 = 0$  :  $y$  has no time trend ( $\varphi_0 = 0$ ), yet  $y$  and the ECE's can have non-zero means and –constants respectively.
- 3)  $\varphi_0 \neq 0, \varphi_1 = 0$  &  $\beta_0 \neq 0, \beta_1 = 0$  :  $y$  has linear time trend ( $\varphi_0 \neq 0$ ), which the EC term removes (in addition to the unit root) – because  $\beta_1 = 0$ .



# VECM: Deterministic specifications



- 4)  $\varphi_0 \neq 0$ ,  $\varphi_1 = 0$  &  $\beta_0 \neq 0$ ,  $\beta_1 \neq 0$ : Here  $y$  follows a time trend, while the ECE also has a time trend component. What this implies is that there is no LR equilibrium, as the series are effectively growing apart at rate  $t$ ...
  - This makes interpretation of the cointegrating parameters more challenging, as it effectively implies that the cointegrating relationships would not be stable over time.
  - This is useful, however, for trend stationary series...
- 5) Typically, the case where  $\varphi_1 \neq 0$  is ignored, as it would imply a quadratic time trend in  $y$ , which is regarded as somewhat unrealistic for most applications in finance.



# Interpreting the results from a VECM

---



- Similar to VARs, the main strength of these models are not to interpret individual coefficient values, but rather to study which variables respond significantly to deviations from the long term relationships, i.e. which adjust to disequilibrium.
- If a variable does not adjust to disequilibrium, it can be interpreted effectively as ***weakly exogenous***.
- Impulse response functions are also useful tools when fitting a VECM. The interpretations are much like that which we established in the VAR form.



# IRFs in a VECM

---



- Refer back to the first semester to see the usefulness of the IRFs in understanding the dynamics of the series in the system.
- Note that the only difference is that many of the impulses would not die away to zero if the series has a unit root.



# Steps to fitting the best VECM

---



- After establishing that the series included are all of the same order & cointegrated – we need to find the level of cointegration ( $r$ ).
  - Thereafter we need to specify the number of lagged differences to include in the VECM.
  - Thereafter we need to specify the form of the VECM in terms of the deterministic factors that we ought to include.
  - This can be done by looking at the data series (qualitative call), or by viewing the AIC and SBIC selection statistics for all the possible combinations of deterministic specifications. This latter option can be easily summarized on the next slide...



# Cointegration



- Highlight all the series
- Select: Quick – cointegration
- Select option 6: Summary of all the sets of assumptions

Johansen Cointegration Test

Cointegration Test Specification

Deterministic trend assumption of test

Assume no deterministic trend in data:

- 1) No intercept or trend in CE or test VAR
- 2) Intercept (no trend) in CE - no intercept in VAR

Allow for linear deterministic trend in data:

- 3) Intercept (no trend) in CE and test VAR
- 4) Intercept and trend in CE - no intercept in VAR

Allow for quadratic deterministic trend in data:

- 5) Intercept and trend in CE - intercept in VAR

Summary:

- 6) Summarize all 5 sets of assumptions

\* Critical values may not be valid with exogenous variables; do not include C or Trend.

Exog variables\*

Lag intervals

14

Lag spec for differenced endogenous

Critical Values

- MHM
- Osterwald-Lenum

Size 0.05

OK Cancel



Sample: 1/02/2002 4/30/2013  
Included observations: 2828  
Series: FINANCIALS HEALTH\_CARE INDUSTRIALS OIL\_AND\_GAS BASIC\_MATE...  
Lags interval: 1 to 4

Selected (0.05 level\*) Number of Cointegrating Relations by Model

| Data Trend: | None                     | None                  | Linear                | Linear             | Quadratic          |
|-------------|--------------------------|-----------------------|-----------------------|--------------------|--------------------|
| Test Type   | No Intercept<br>No Trend | Intercept<br>No Trend | Intercept<br>No Trend | Intercept<br>Trend | Intercept<br>Trend |
| Trace       | 2                        | 3                     | 5                     | 2                  | 2                  |
| Max-Eig     | 2                        | 2                     | 2                     | 2                  | 2                  |

\*Critical values based on MacKinnon-Haug-Michelis (1999)

- The first part of the summary output gives us the best estimation of the rank. In this case,  $n = 5$ , and the rank  $r = 2$ . As  $r > 0$  &  $< 5$ , it suggests a stable VECM can be estimated. Now scroll down and we get an idea of the deterministic form required for the VECM...



- After finding the best rank on the previous slide (which was rank 1), we then need to find the best deterministic specification below:

The best specification would then be the one where the AIC and SBIC is minimized and Log Likelihood maximized, for the optimal cointegration level (in this case 2).

The output suggests the second last column...

Information Criteria by Rank and Model

| Data Trend:  | None                  | None               | Linear             | Linear          | Quadratic       |
|--|-----------------------|--------------------|--------------------|-----------------|-----------------|
| Rank or No. of CEs   | No Intercept No Trend | Intercept No Trend | Intercept No Trend | Intercept Trend | Intercept Trend |
| Log Likelihood by Rank (rows) and Model (columns)              |                       |                    |                    |                 |                 |
| 0  | -67812.98             | -67812.98          | -67803.21          | -67803.21       | -67792.90       |
| 1  | -67784.67             | -67784.43          | -67781.75          | -67779.47       | -67772.55       |
| 2  | -67770.92             | -67765.01          | -67762.75          | -67759.78       | -67755.73       |
| 3  | -67764.37             | -67754.81          | -67754.77          | -67749.67       | -67745.77       |
| 4  | -67759.16             | -67749.36          | -67749.36          | -67743.25       | -67741.03       |
| 5  | -67758.79             | -67746.72          | -67746.72          | -67740.14       | -67740.14       |
| Akaike Information Criteria by Rank (rows) and Model (columns) |                       |                    |                    |                 |                 |
| 0  | 48.02898              | 48.02898           | 48.02561           | 48.02561        | 48.02185        |
| 1  | 48.01603              | 48.01657           | 48.01750           | 48.01660        | 48.01454        |
| 2  | 48.01338              | 48.01062           | 48.01114           | 48.01045        | 48.00971*       |
| 3  | 48.01582              | 48.01118           | 48.01257           | 48.01108        | 48.00974        |
| 4  | 48.01921              | 48.01511           | 48.01581           | 48.01432        | 48.01346        |
| 5  | 48.02602              | 48.02102           | 48.02102           | 48.01990        | 48.01990        |
| Schwarz Criteria by Rank (rows) and Model (columns)            |                       |                    |                    |                 |                 |
| 0  | 48.23929*             | 48.23929*          | 48.24642           | 48.24642        | 48.25318        |
| 1  | 48.24736              | 48.25001           | 48.25935           | 48.26055        | 48.26690        |
| 2  | 48.26574              | 48.26718           | 48.27402           | 48.27753        | 48.28310        |
| 3  | 48.28310              | 48.28310           | 48.28310           | 48.28310        | 48.28310        |
| 4  | 48.28310              | 48.28310           | 48.28310           | 48.28310        | 48.28310        |
| 5  | 48.28310              | 48.28310           | 48.28310           | 48.28310        | 48.28310        |

**But** we typically don't fit quadratic trends on financial time series data, so we choose the next best – form 4:  $\varphi_0 \neq 0$ ,  $\varphi_1 = 0$  &  $\beta_0 \neq 0$ ,  $\beta_1 \neq 0$



UNIVERSITEIT•STELLENBOSCH•UNIVERSITY  
jou kennisvennoot • your knowledge partner

# Differing Orders of Integration



---



Department of Economics

DEPARTMENT OF  
**ECONOMICS**



# Integration of differing orders

---



Suppose we have the following problem:

- We have a  $I(1)$  variable and also a  $I(0)$  variable, or any combination of series that do not share the same level of integration.
- How do we treat these series?
- Can we pile them together to estimate our model?



# Integration of differing orders

---



- Suppose we are interested in a bivariate relationship having the following integration levels:
- $I(0)$  &  $I(0)$ : Standard asymptotic arguments hold.
- $I(1)$  &  $I(0)$ : This is an ***unbalanced*** regression. Here the  $I(0)$  cannot explain the LR variation in  $I(1)$ , as the latter process is driven by a unit root process (and vice versa). **We then need to difference the  $I(1)$  series and see whether its short run dynamics can be explained by / explain the  $I(0)$  series.**



# Integration of differing orders

---



- $I(1)$  &  $I(1)$ : We now either have cointegration or spurious regressions.
  - Spurious regressions are models where we have two independent  $I(1)$  processes. This leads to  $\beta$  estimates that can effectively take any value, and typically show highly significant *t* – stats.
  - The non-sensical results can therefore not serve as having any sound economic interpretation value.
  - When regressing two  $I(1)$  variables, you should at least make sure that the residuals are  $I(0)$  – **which (by definition) is the same as establishing a cointegrating relationship!** 😊



# Encountering problems with high persistence that may not be $I(1)$ ...



- Some authors tend to take differences (or use cointegration testing & VECMs) on series that show **strong persistence**, yet does not reject ADF...
- The authors then motivate why they take the differences.
- This problem can also be seen when we regress, e.g., share price returns (Dlogged closing prices), on a highly persistent predictor (such as Dividend Yield). Caution then needs to be taken when doing this, as the strong persistence might lead to non-rejection of hypotheses tests... and might warrant differencing the highly persistent series (as shown by its  $AR(1)$  term, e.g.)



# The END!

---



- Let's now go and apply it in EVIEWS.