Applied Econometrics

Regression with Time Series Variables

Jaromír Baxa

Introduction

- Univariate stationary models: good for descriptive analysis and for forecasting, especially short-term forecasting (one period ahead).
- However, many different things can be done with time series!
    - Forecasting at longer horizons
    - Co-movement of variables
    - Measurement of the effect of one variable to another.
- With time series data, such analysis is sometimes complicated by serial correlation, and because variables affect each other often with a time lag.
- Various approaches:
- Single equation models (Autoregressive distributed lags model - ARDL, Error correction model – ECM, ...)
- Multiple equation models (VAR and VECM); with many extensions.
- This lecture: single equation models.

Outline

- Time series regression when X and Y are stationary
- Nonstationary variables I: Spurious regression problem
- Nonstationary variables II: Cointegration
- The error-correction model
- Time series regression when X and Y are non-stationary and non-cointegrated
- Summary

Time series regression when X and Y are stationary

- Simple OLS $Y_t = a + bX_t + u_t$ does not exploit the time dimension.
- Therefore, let's start with an autoregressive distributed lag model; ARDL(1,1):

$$Y_t = \alpha + \rho_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$

- First and foremost, the estimation of multivariate models depends on whether the series used in the analysis (Y_t, X_t) are stationary or not.
- Additionally, it is assumed, that both variables in regression are of the same order
    - Stationary variable X_t cannot successfully explain Y_t with unit root.
- In practice, we should pre-test all variables for the existence of unit root prior further analysis, and the baseline ARDL models shall be estimated only with stationary variables.

Time series regression when X and Y are stationary

- In general, the ARDL models can have many lags, the order is denoted as (p,q). Also, infinite-lag modifications exist.

$$Y_t = \alpha + \rho_1 Y_{t-1} + \ldots + \rho_p Y_{t-p} + \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_q X_{t-q} + \varepsilon_t$$

- If all variables are stationary, estimation of the ARDL model (or of regression model with time series variables and some arbitrary lags) is easy.
- Lag length usually selected automatically using information criteria.
- The OLS works fine and standard tests such as *t*-tests and *F*-tests can be carried out to select significant and insignificant variables and to decide about the lag length.
- The ARDL coefficients are used to calculate multipliers:
  - The impact multiplier: $\dfrac{\partial Y_t}{\partial X_t} = \dfrac{\beta_0}{1 - \Sigma_{i=1}^{p} \rho_i}$
  - The long-run multiplier $\dfrac{\partial Y_{t+\infty}}{\partial X_t} = \dfrac{\Sigma_{j=0}^{q} \beta_j}{1 - \Sigma_{i=1}^{p} \rho_i}$

Nonstationary variables I: Spurious regression problem

- In many cases, we study behavior of time-series that are non-stationary.
- ***Spurious regression problem*** might arise: if both variables are growing, the OLS finds a strong relationship although the variables do not have anything in common.
- Suppose a model

$$Y_t = \alpha + \beta X_t + u_t$$

- Even if the true value of the coefficient $\beta=0$, the OLS estimate can be non-zero, R^2 can be high and the p-values of test statistics can be low as well.

Nonstationary variables I: Spurious regression problem

- In many cases, we study behavior of time-series that are non-stationary.
- ***Spurious regression problem*** might arise: if both variables are growing, the OLS finds a strong relationship although the variables do not have anything in common.
- Suppose a model

$$Y_t = \alpha + \beta X_t + u_t$$

- Even if the true value of the coefficient $\beta=0$, the OLS estimate can be non-zero, R^2 can be high and the p-values of test statistics can be low as well.

➔ ***Consequences of spurious regression***

  – Slope coefficient β biased.

  – The t-statistics are increasing with the size of the sample

  – The R^2 converges to 1 with $T \rightarrow \infty$ => *superconsistency*

➔ ***Therefore, you shall be cautious when running a regression of Y on X when both variables have unit roots.***

Nonstationary variables I: Spurious regression problem

- Example: Variables correlated just by coincident
- Two independent RW processes of the same structure
  - OLS regression => positive and significant correlation likely, perhaps DW statistics will suggest autocorrelation in residuals
  - OLS regression on first differences => likely not any significant relationship

```
set.seed(458)
e1 = rnorm(250)
e2 = rnorm(250)
y1 = cumsum(e1)
y2 = cumsum(e2)
summary(OLS(y1~y2))
Coefficients Value   Pr(>|t|)
(Intercept)  6.7445  0.0000
Y2           0.4083  0.0000

Regression Diagnostics:
R-Squared 0.2066
Adjusted R-Squared 0.2034
Durbin-Watson Stat 0.0328
```

```
summary(OLS(diff(y1)~diff(y2)))
Coefficients Value   Pr(>|t|)
(Intercept) -0.0565  0.3991
diff(y2)     0.0275  0.6683

R-Squared 0.0007
Adjusted R-Squared -0.0033
Durbin-Watson Stat 1.9356
```

Nonstationary variables I: Spurious regression problem

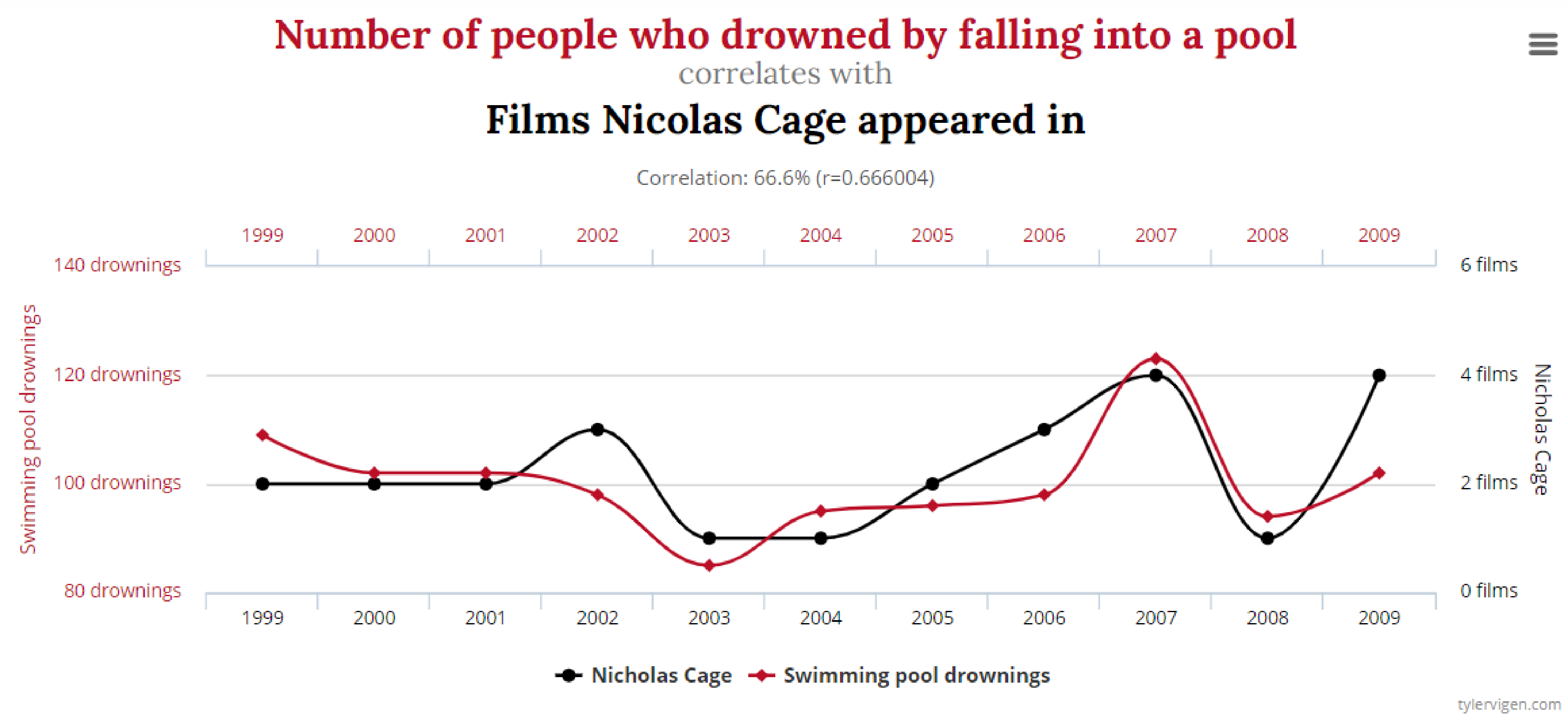
- Another reason for spurious regression: Variables are correlated but not causally related due to a presence of a third confounding variable that affects both X and Y.
- Examples:

Nonstationary variables I: Spurious regression problem

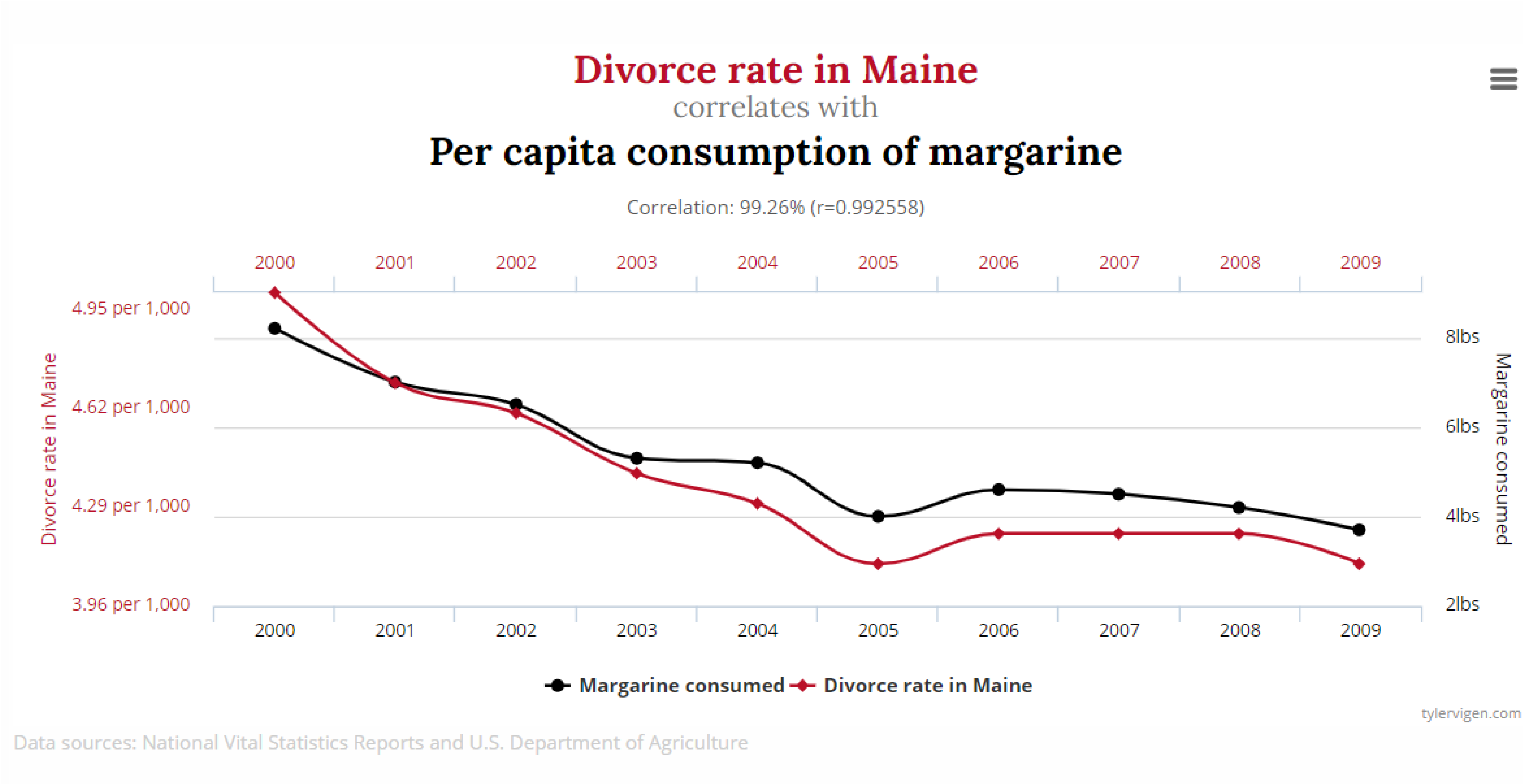
- Another reason for spurious regression: Variables are correlated but not causally related due to a presence of a third confounding variable that affects both X and Y.
- Examples:
    - Ice cream sales and drownings in swimming pools => likely driven by good weather

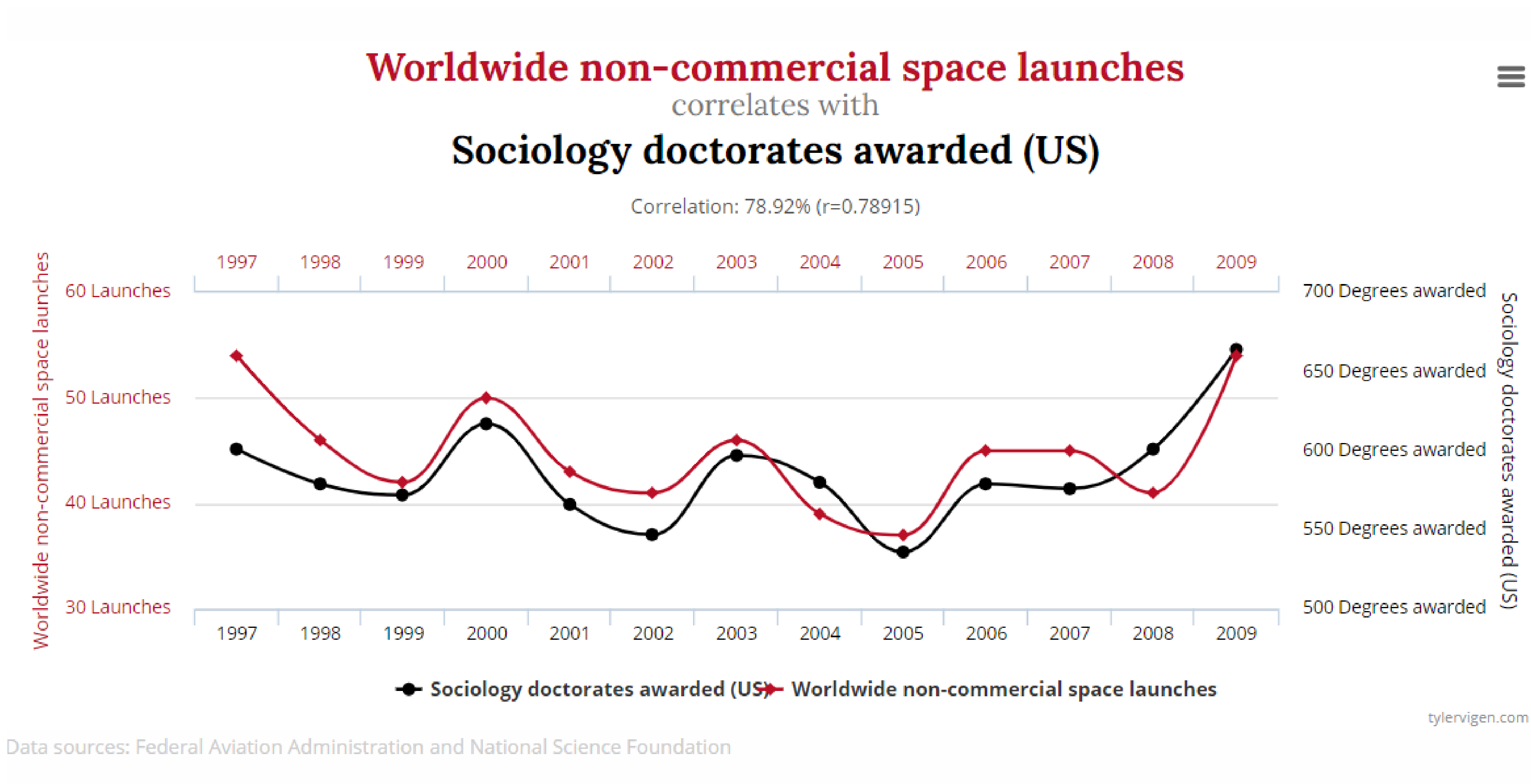
Nonstationary variables I: Spurious regression problem

- Another reason for spurious regression: Variables are correlated but not causally related due to a presence of a third confounding variable that affects both X and Y.
- Examples:
  - Ice cream sales and drownings in swimming pools => likely driven by good weather
  - Nominal macro variables => driven by price level (wages, industrial production or productivity).



Data sources: Centers for Disease Control & Prevention and Internet Movie Database


Divorce rate in Maine
correlates with
Per capita consumption of margarine
Correlation: 99.26% (r=0.992558)
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
4.95 per 1,000
4.62 per 1,000
4.29 per 1,000
3.96 per 1,000
Divorce rate in Maine
8lbs
6lbs
4lbs
2lbs
Margarine consumed
Margarine consumed
Divorce rate in Maine
tylervigen.com
Data sources: National Vital Statistics Reports and U.S. Department of Agriculture



Worldwide non-commercial space launches
correlates with
Sociology doctorates awarded (US)
Correlation: 78.92% (r=0.78915)
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
60 Launches
50 Launches
40 Launches
30 Launches
Worldwide non-commercial space launches
700 Degrees awarded
650 Degrees awarded
600 Degrees awarded
550 Degrees awarded
500 Degrees awarded
Sociology doctorates awarded (US)
Sociology doctorates awarded (US)
Worldwide non-commercial space launches
tylervigen.com
Data sources: Federal Aviation Administration and National Science Foundation

Nonstationary variables I: Spurious regression problem

How to avoid spurious regression?

1) Avoid data mining (or use the outcomes with care).
2) Use randomized controlled trials or natural experiments to derive causal relationships (not always feasible)
3) Be careful about regressions with nonstationary variables and use a methodology that can account for cointegration.

Nonstationary variables II: Cointegration

- Let's return back to a two-variable regression model with variables of order I(1):

$$Y_t = \alpha + \beta X_t + u_t$$

- Assume that their long-term relationship is stable and systemic. Then, if the residuals $e_t = Y_t - \alpha - \beta X_t$ are $I(0)$, we call the pair of series X_t and Y_t as being **cointegrated**.

Cointegration and Error Correction

Why is the existence of cointegration interesting?

- It does not arise often: a linear combination of two I(1) series is in general I(1) as well. But sometimes, the unit roots in both variables 'cancel each other out'.
- If the time series are cointegrated, the true cointegrating linear combination defines a long run equilibrium. The existence of the equilibtium is usually supported by economic theory => estimation of cointegration can be used to estimate and evaluate the very existence of equilibrium relationships.
- Equilibrium: $\beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_n x_{nt} = e_t$ with $E(e_t) = 0$
- The economic variables are not exactly at the equilibrium, so the equilibrium error e_t appears. Nevertheless, they never drift far away as they are always forced back towards it.
- Furthermore, cointegration implies error-correction; that is, the movement of a variable is function of its past deviation from the equilibrium, e_{t-1} (Engle and Granger, 1987).

Formal definition

A (n×1) vector time series x_t *consisting of I(1) series is said to be cointegrated if there exist any non-zero vector* β *such that a linear combination* $\beta' x_t$ *is stationary I(0).*

- The β is refered as *cointegrating vector*.
- Cointegrating vector is not unique – if β, then $\lambda\beta$ as well => β usually normalized to $|\beta|=1$.

Cointegration and Correlation

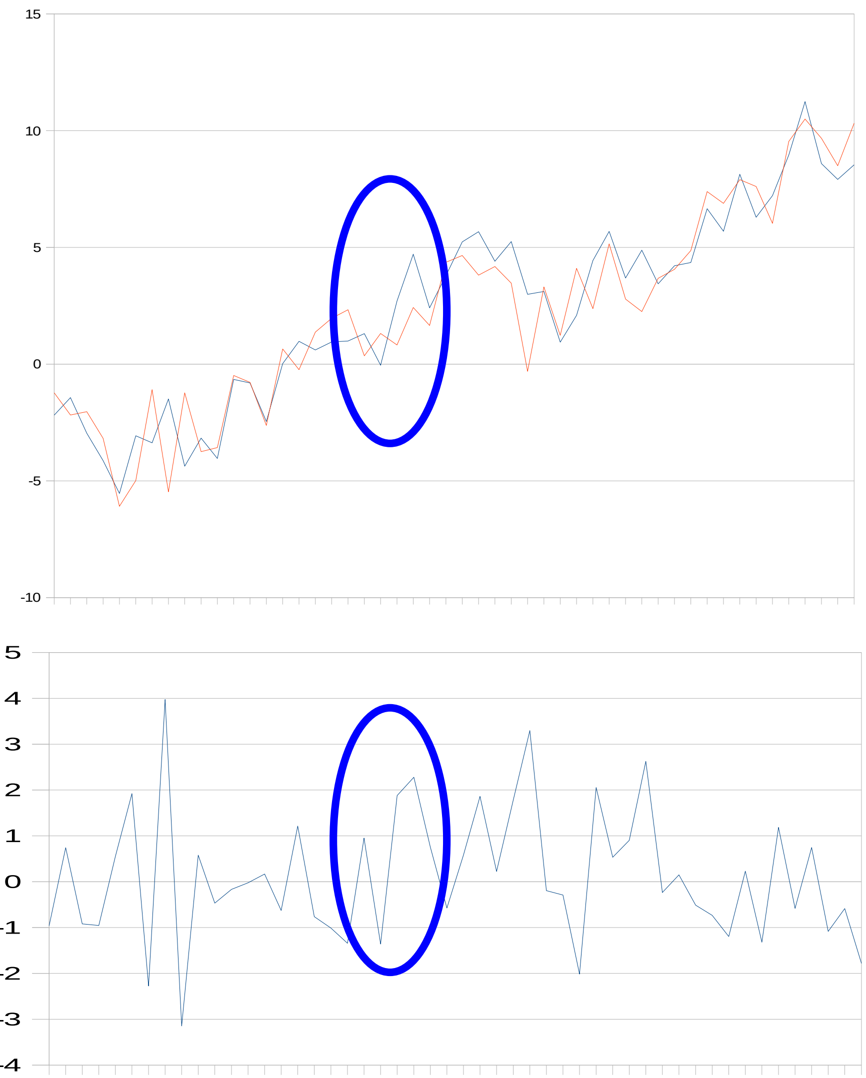
Correlation – if one variable moves up, probably the second does the same.

Cointegration – if one variable moves up, the second either does the same or the first decreases after some time to keep their long-term relationship stable.

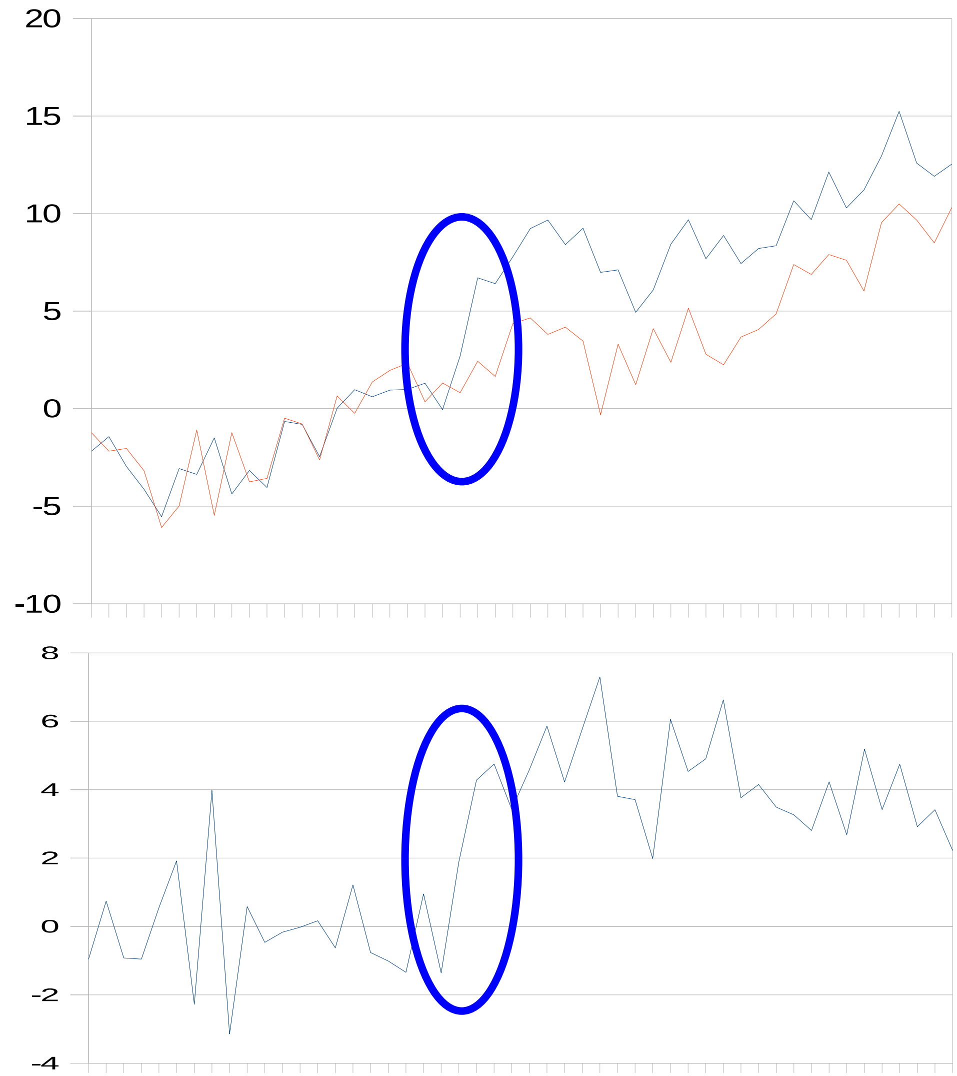
In fact it means that the two (or more) variables cannot wander off in log term and the deviations from their long-term stable relationship are only temporary.

=> Cointegration indicates more tight relationship between variables than correlation does.

Cointegration and Correlation: An Illustration



Shock to „Blue“ caused only temporary shift from the equilibrium relation; corr = 0.941



Shock to „Blue“ caused shift in equilibrium; corr = 0.923

Examples

Example 1

- Consumption, output and the permanent income hypothesis:

➔ $$c_t = c_t^P + c_t^T$$

- Since the permanent consumption c^P is proportional to permanent income, we can write:

➔ $$c_t = \beta\, y_t^P + c_t^T$$

- If the permanent income hypothesis holds, the transitory consumption shall be a linear combination of two $I(1)$ variables given by $c_t - \beta\, y^P$, and it should be stationary.

Example 2

- Relation among money and real economy: M = kPY => Demand of money
- Take logs, assume k being a function of the interest rate:

$$m_t = \beta_0 + \beta_1 p_t + \beta_2 y_t + \beta_3 r_t + e_t$$

- Solving for e_t:

$$e_t = m_t - \beta_0 - \beta_1 p_t - \beta_2 y_t - \beta_3 r_t$$

➔ Linear combination of m_t, y_t, r_t, p_t is supposed to be stationary.

Example 3

- Law of one price: The prices of goods expressed in common currency should be identical, so the relationship $S_t{*}P_{t;foreign} = P_t$ should hold at least in the long-term.
- Deviations of this parity should be temporary (due to arbitrage) => that is, linear combination of them should be stationary.

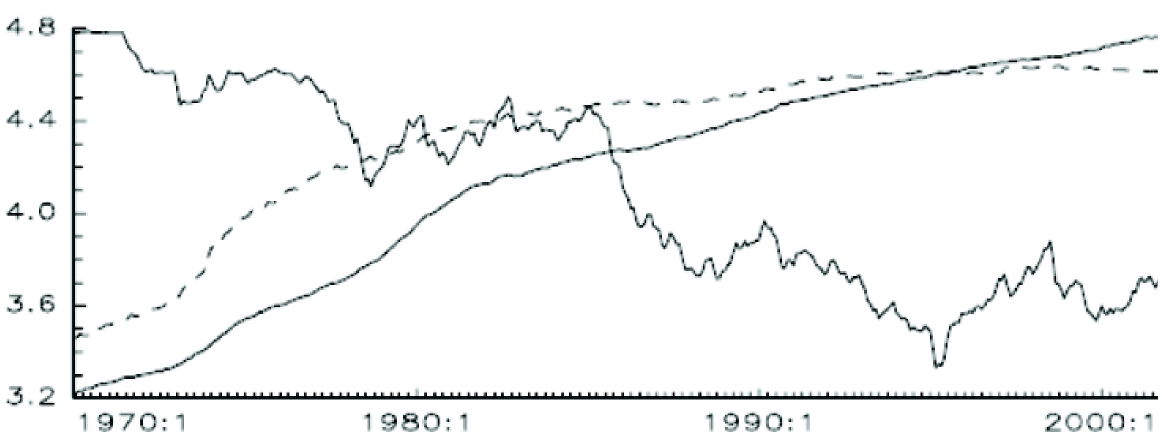


Figure 1.1: Logarithm (rescaled) of the Japanese yen/US dollar exchange rat creasing solid line), logarithm of seasonally adjusted US consumer price index (in ing solid line) and logarithm of seasonally adjusted Japanese consumer price (increasing dashed line), 1970:1 — 2003:5, monthly observations

Figure: Regression residuals

Other examples

Models of equilibrium exchange rates (exchange rate is supposed to be determined by its fundamentals, in particular by net foreign assets, terms of trade and a differential in productivity growth).

Growth theory: GDP, C, I

Covered interest rate parity: forward and spot exchange rates

Fisher equation: nominal interest rates and inflation

Term structure of interest rates and the expectations hypothesis: nominal interest rates at different maturities

Finance – high-frequency or low-frequency relationships: price of assests at different markets; spot and futures prices ; low-frequency: asset prices and fundamentals.

Testing cointegration

- First, we introduce the ***error-correction model*** (ECM), as for any set of I(1) variables, error correction and cointegration are the equivalent representations (the Granger representation theorem).
- ECM: dynamic model, movement of variable related to previous period's gap to long-run trend.
- Assume the model $X_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$

  with X_t and Y_t both $\sim I(1)$
- Subtract X_{t-1} from both sides of equation and get

$$X_t - X_{t-1} = \alpha_0 + (\alpha_1 X_{t-1} - X_{t-1}) + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$$

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$$

Testing cointegration

- First, we introduce the ***error-correction model*** (ECM), as for any set of I(1) variables, error correction and cointegration are the equivalent representations (the Granger representation theorem).
- ECM: dynamic model, movement of variable related to previous period's gap to long-run trend.
- Assume the model $X_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$

  with X_t and Y_t both $\sim I(1)$
- Subtract X_{t-1} from both sides of equation and get

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$$

- Now add: $- \beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and rearrange:

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 Y_t - \beta_0 Y_{t-1} + \beta_0 Y_{t-1} + \beta_1 Y_{t-1} + \varepsilon_t$$

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + \varepsilon_t$$

Testing cointegration

- First, we introduce the ***error-correction model*** (ECM), as for any set of I(1) variables, error correction and cointegration are the equivalent representations (the Granger representation theorem).
- ECM: dynamic model, movement of variable related to previous period's gap to long-run trend.
- Assume the model $X_t = \alpha_0 + \alpha_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon$

  with X_t and Y_t both $\sim I(1)$
- Subtract X_{t-1} from both sides of equation and get

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 Y_t + \beta_1 Y_{t-1} + \varepsilon_t$$

- Now add: $- \beta_0 Y_{t-1} + \beta_0 Y_{t-1}$ and rearrange:

$$\Delta X_t = \alpha_0 + \rho_1 X_{t-1} + \beta_0 \Delta Y_t + \theta_1 Y_{t-1} + \varepsilon_t$$

- The ECM is then:

$$\Delta X_t = \alpha_0 + \rho_1 (X_{t-1} - \gamma Y_{t-1}) + \beta_0 \Delta Y_t + \varepsilon_t$$

(Note that when LHS is I(0), ΔY_t is $I(0)$ and if X_t and Y_t cointegrated, e_t must be $I(0)$, too.)

Engle-Granger Procedure

To test the cointegration, we follow the Engle-Granger procedure:

1)Test the order of integration for all variables by ADF

2) Estimate (by OLS) $X_t = \alpha_0 + \beta_0 Y_t + e_t$, where Y_t is vector of variables, save the residuals e_t (the e_t is a candidate for the error correction term).

3) The residuals should be $I(0)$, ADF test with the critical values from Engle-Yoo (1987) is used.

4) Estimate the error-correction model $\Delta X_t = \alpha_0 + \beta \Delta Y_t + \rho e_{t-1} + u_t$,
(sometimes lags of ΔX_t and ΔY_t needed; e_{t-1} comes from the step 2)

5) Evaluate the model adequacy (the parameter ρ is significant and the value is that the magnitude of an error is diminishing over time. Then the parameter ρ can be interpreted as the speed of adjustment as the e_{t-1} is the error correction term.)

Drawbacks of the Engle-Granger approach

- Low power: cointegration often rejected even when it is present.
- It can be extended for more than two variables, however in this case there can be more than one cointegrating relationships – and this situation cannot be treated using this approach

Alternatives to the Engle-Granger approach

- Single equation models
  - **Dynamic OLS (DOLS)**
  - **Autoregressive distributed lags (ARDL)**
  - Fully modified OLS (FM-OLS)
  - Canonical cointegrating regression (CCR)
- Multi-equation models (VECM)

Alternatives to the Engle-Granger approach

- Single equation models
  - **Dynamic OLS (DOLS)** – follows Stock and Watson (1993)

$$y_t = c + \beta x_t + \sum_{i=-p}^{q} \Delta x_t + \varepsilon_t$$

    - Contains leads and lags of first differences of the independent variable
    - Number of leads and lags selected using information criteria
  - Autoregressive distributed lags (ARDL)
  - Fully modified OLS (FM-OLS)
  - Cannonical cointegrating regression (CCR)
- Multi-equation models (VECM)

Alternatives to the Engle-Granger approach

- Single equation models
    - Dynamic OLS (DOLS)
    - **Autoregressive distributed lags (ARDL)**
        - Most versatile, even for mixture of I(0) and I(1) variables with some of them cointegrated.
        - ARDL / Bounds Testing methodology by Pesaran and Shin (1999)
        - $\Delta y_t = \beta_0 + \Sigma\, \beta_i \Delta y_{t-i} + \Sigma \gamma_j \Delta x_{1t-j} + \Sigma \delta_k \Delta x_{2t-k} + \varphi e_{t-1} + \varepsilon_t$

          with z_t being the error correction term, i.e., residuals from $y_t = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + e_t$
        - $\Delta y_t = \beta_0 + \Sigma\, \beta_i \Delta y_{t-i} + \Sigma \gamma_j \Delta x_{1t-j} + \Sigma \delta_k \Delta x_{2t-k} + \theta_0 y_{t-1} + \theta_1 x_{1t-1} + \theta_2\, x_{2t-1} + \varepsilon_t$
        - Make sure the residuals are white noise, then F-test on H_0: $\theta_0 = \theta_1 = \theta_2 = 0$. When rejected, cointegration present. Again, non-standard critical values.

Time series regression when X and Y are non-stationary and non-cointegrated

- Use ARDL or OLS with variables in first differences. That’s it :-)

Summary

- We have introduced main principles of multivariate time series analysis within the single equation framework.
- We need to care about stationarity.
  - If all variables I(0), ARDL works fine.
- When working with I(1) variables, there is a risk of spurious regression that cannot be identified using statistical tests.
- I(1) variables could be cointegrated. Cointegration is a strong, long-term relationship among variables. It implies much stronger mutual dependence than correlation.
- It occurs if variables share a common trend or if there is some form of equilibrium relationship – as in money demand equation.
  - If the I(1) variables are not cointegrated, one should utilize first differences in regression models (ARDL).
  - If variables are I(1) and cointegrated, OLS in levels (to get long-run coefficients – but not s.e.'s!), Error-correction model, DOLS or ARDL with error correction term.
- **In all cases, the researcher must be aware of the structure of the model.**

Readings

- This lecture was based on Koop, G. (2008) Introduction to Econometrics, Ch. 7 with some extensions.

Further reading:

- The Royal Swedish Academy of Sciences (2003): Time Series Econometrics: Cointegration and Autoregressive Conditional Heteroscedasticity, downloadable from:
http://www-stat.wharton.upenn.edu/~steele/HoldingPen/NobelPrizeInfo.pdf

- Granger, C. W.J. (2003): Time Series, Cointegration and Applications, Nobel lecture, December 8, 2003, downloadable from: http://ideas.repec.org/p/cdl/ucsdec/2004-02.html

- Almost all textbooks cover the introduction to cointegration: Either to Engle-Granger procedure (single equation procedure), or to the Johansen multivariate framework. See for example Enders, W.: Applied Econometric Time Series (2nd edition).

- Moosa, Imad A. (2017): Blaming suicide on NASA and divorce on margarine: The hazard of using cointegration to derive inference on spurious correlation. Applied Economics, 49, 2017.

On ARDL:

- http://davegiles.blogspot.cz/2013/06/ardl-models-part-ii-bounds-tests.html

- ARDL with R packages `dynamac, ardl`
https://cran.r-project.org/web/packages/dynamac/vignettes/dynamac-vignette.html
https://search.r-project.org/CRAN/refmans/ARDL/html/ardl.html
https://search.r-project.org/CRAN/refmans/ARDL/html/multipliers.html