Cointegration

(Hamilton: Chapters 19, Zivot and Wang: Chapter 12)

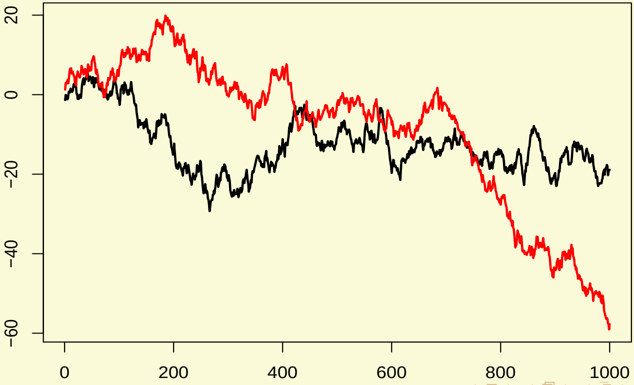
Isabel Casas

SYDDANSK UNIVERSITET

- Spurious regression
- Cointegration
  - Phillip's triangular representation
  - Error-correction representation
- Test for cointegration
  - When the cointegrating vector is pre-defined
  - Estimating the cointegrating vector

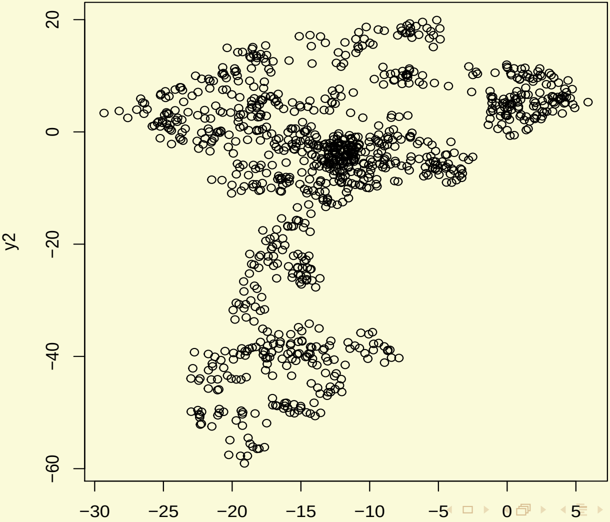
Example: spurious regression

```
> epsilon1=rnorm(1000)
> epsilon2=rnorm(1000)
> y1=cumsum(epsilon1)
> y2=cumsum(epsilon2)
```



Example: spurious regression

Are they linearly related?



Example: spurious regression

Run the regression: $y_{1t} = \beta y_{2t} + e_t$

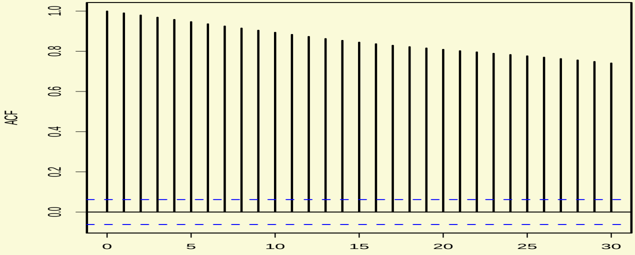
```
> summary(model)

Call:
lm(formula = y1 ~ -1 + y2)

Residuals:
    Min      1Q  Median      3Q     Max
-30.539 -14.433  -9.310  -2.760   6.783

Coefficients:
   Estimate Std. Error t value Pr(>|t|)
y2  0.36370    0.01898   19.17   <2e-16 ***
---
Signif. codes:  0 âĂŸ***âĂŹ 0.001 âĂŸ**âĂŹ 0.01 âĂŸ*âĂŹ 0.05 âĂŸ.âĂŹ 0.1 âĂŸ âĂ

Residual standard error: 12.41 on 999 degrees of freedom
Multiple R-squared: 0.2689,      Adjusted R-squared: 0.2681
F-statistic: 367.4 on 1 and 999 DF,  p-value: < 2.2e-16
```

Syddansk Universitet

Example: spurious regression

$$y_{1t} = 0.36 y_{2t} + e_t \qquad \hat{\sigma} = 12.41$$

- The slope coefficient is 0.36 and appears as highly significant
- The $R^2 = 0.27$ is quite good and it gets better as the sample size increases
- The residuals autocorrelation:



Example: spurious regression

The Durbin-Watson test

$$H_0: \text{ autocorrelation of the innovations is zero}$$

```
> library(lmtest)
> dwtest(model)

        Durbin-Watson test

data:  model
DW = 0.0068, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater tha
```

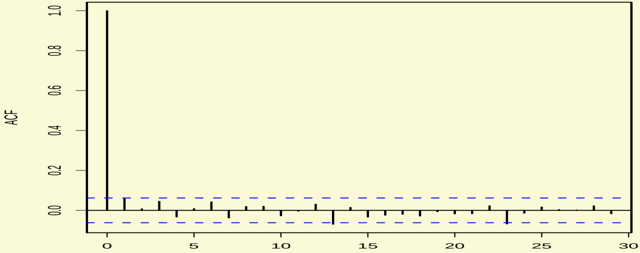
- These are the properties of a spurious regression
- It happens because the two variables are I(1) and no *cointegrated*
- If we regress Δy_{1t} on Δy_{2t}, the correct relationship is revealed

Example: spurious regression

```
> deltay1=diff(y1)
> deltay2=diff(y2)
> model2<-lm(deltay1~-1 + deltay2)

Call:
lm(formula = deltay1 ~ -1 + deltay2)

Residuals:
     Min       1Q   Median       3Q      Max
-3.07848 -0.72668  0.00278  0.66122  3.06710

Coefficients:
        Estimate Std. Error t value Pr(>|t|)
deltay2 0.006419   0.030439   0.211    0.833

Residual standard error: 0.957 on 998 degrees of freedom
Multiple R-squared: 4.456e-05,       Adjusted R-squared: -0.0009574
F-statistic: 0.04447 on 1 and 998 DF,  p-value: 0.833
```

Example: spurious regression

- The residuals of the regression of the first difference are now serially uncorrelated
- However, we have lost information taking the first difference
- And, it might be difficult to interpret the results back on the real series y_{1t}, y_{2t}



Statistical implication of spurious regression

$$y_{1t} = \beta y_{2t} + e_t$$

- y_{1t} and y_{2t} are I(1) variables
- Since they are not cointegrated, the true value of $\beta = 0$
- $\hat{\beta}$ do not converge in probability to zero but instead converges in distribution to a non-normal random variable not necessarily centered at zero.
- The usual OLS t-statistics for testing $\beta = 0$ diverge to $\pm\infty$ as $T \rightarrow 0$
- The usual R^2 converges to 1 as $T \rightarrow \infty$
- Regression with I(1) variables only makes sense if they are cointegrated.

SYDDANSK UNIVERSITET

Applications where this is important

Examples of financial series that depend of each other.

Definition: cointegration

A vector time series $\mathbf{y}_t = \{y_{1t}, \ldots, y_{nt}\}'$ is said to be *cointegrated of order 1* if

1. Each of the series taken individually are I(1) (they have an unit root) and,
2. There is a combination of the series $\mathbf{a'y_t}$ which is I(0) for some non zero $\mathbf{a}$

- Although changes in the individual elements of $\mathbf{y}_t$ drift apart, there is a long-run equilibrium relation trying to keep these components together (they cannot escape from each other)
- $\mathbf{a}$ is called the *cointegrating vector*
- Simile: A drunk and her dog (Murray, 1994)

Q: cointegrated of order n?

Definition: cointegration

- This linear relationship $\mathbf{a}$ is also called the *atractor* or as Murray says an *error-correction mechanism*
- The two variables are allowed to diverge in the short-run,
- In the long-run they have to converge to a common region

Definition: cointegration

- $y_{1t} \sim I(0) \Rightarrow a + by_{1t} \sim I(0)$
- $y_{1t} \sim I(1) \Rightarrow a + by_{1t} \sim I(1)$
- $y_{1t}, y_{2t} \sim I(0) \Rightarrow ay_{1t} + by_{2t} \sim I(0)$
- $y_{1t} \sim I(0), y_{2t} \sim I(1) \Rightarrow ay_{1t} + by_{2t} \sim I(1)$
- $$y_{1t}, y_{2t} \sim I(1) \Rightarrow \begin{cases} ay_{1t} + by_{2t} \sim I(1) & \text{in general} \\ ay_{1t} + by_{2t} \sim I(0) & \text{if they are cointegrated} \end{cases}$$

Example: cointegration

Theory of purchasing power parity (PPP): "Apart from transportation costs, goods should sell for the same effective price in two countries"

$$P_t = S_t P_t^*$$

- P_t is the price in USA (US \$)
- P_t^* is the price in Denmark (DKK)
- S_t is the exchange rate

Taking the log

$$p_t = s_t + p_t^*$$

Example: cointegration

Usually that equation does not hold: measurement errors, transportion costs and differences in quality (z_t)

$$p_t = s_t + p_t^* + z_t$$

A weaker theory says that $z_t \sim I(0)$ even though $p_t, s_t, p_t^* \sim I(1)$. These means that they are cointegrated.

VAR in cointegration

Example of cointegrated system:

$$y_{1t} = \gamma y_{2t} + u_{1t}$$
$$y_{2t} = y_{2,t-1} + u_{2t}$$

Can this be written as a VAR for the differenced data?

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} 1-L & \gamma L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

Find the mistake in the equation above.

VAR in cointegration

However, we can rewrite the VAR in levels as:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -1 & \gamma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} + \gamma u_{2t} \\ u_{2t} \end{bmatrix}$$

With cointegrated systems, we have to include lagged levels along with lagged differences to explain $\mathbf{\Delta y}_t$

Cointegrating vector

Let $\mathbf{y_t} = \{y_{1t}, y_{2t}\}'$ be a cointegrated vector of order one which components are composed by a common I(1) part and different I(0) parts.

$$\mathbf{y_t} = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = (a, b) \begin{pmatrix} w_t \\ w_t \end{pmatrix} + \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix}$$

$$\begin{aligned} y_{1t} =& a w_t + \tilde{y}_{1t} \qquad & w_t \sim I(1),\ \tilde{y}_{1t} \sim I(0) \\ y_{2t} =& b w_t + \tilde{y}_{2t} \qquad & \tilde{y}_{2t} \sim I(0) \end{aligned}$$

Then, what is

- $3(a, b)\mathbf{y_t}$?

Cointegrating vector

- So there are infinite number of cointegrating vectors.
- We find the normalised one, the one that has a 1 as first element
- In out example above $(1, a/b)$
- If $\overbrace{\mathbf{y_t}}^{n\times 1} = (\mathbf{y_{1,t}}, \ldots, \mathbf{y_{nt}})$ how many cointegration vectors can we have?
- There might be $h < n$ cointegrating vectors $(n \times 1)$ $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_h}$ such that $\mathbf{a}_i' y_t \sim I(0)$
- This means that there are $n - h$ *common I(1) stochastic trends*
- All $\mathbf{a_i}$ linearly independent (there is no scalar b such that $\mathbf{a_i} = b\mathbf{a_j}$ for $i \neq j$)

Cointegrating vectors

- $\mathbf{A'y_t} \sim I(0)$ vector

$$\underbrace{\mathbf{A}'}_{h\times n} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_h' \end{bmatrix}$$

- These vectors $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_h}$ are not unique,
- So if $\mathbf{A'y_t} \sim I(0)$, then $\mathbf{b'A'y_t} \sim I(0)$ for $\underbrace{\mathbf{b}'}_{1\times h}$ any nonzero vector.
- Therefore the vector $\boldsymbol{\pi} = \mathbf{b'A'}$ could be considered as a cointegrating vector

Syddansk Universitet

Simulating Cointegrated Systems

These systems can be simulated using the Phillip's triangular representation (Phillips, 1991)

Bivariate: 1 cointegrated vector, 1 common trend

- Assume $\mathbf{y_t} = (y_{1t}, y_{2t})'$ is cointegrated,
- and $\mathbf{a} = (1, -a_2)'$ is the cointegrating vector

$$
\begin{aligned}
y_{1t} &= a_2 y_{2t} + e_{1t} & \qquad e_{1t} \sim I(0) \\
y_{2t} &= y_{2,t-1} + e_{2t} & \qquad e_{2t} \sim I(0)
\end{aligned}
$$

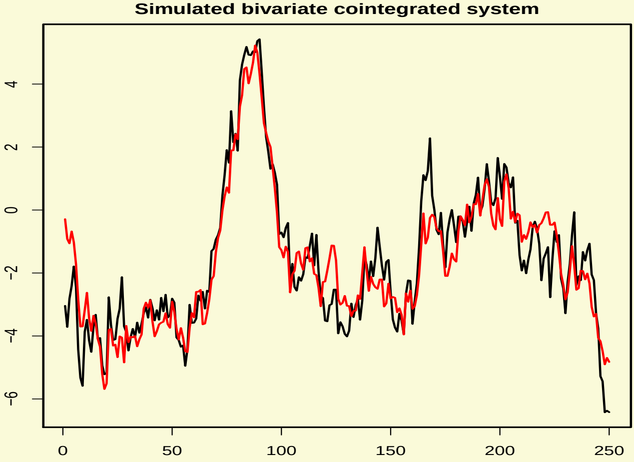
- First eq. describes the long-run equilibrium relationship with I(0) disequilibrium error e_{1t}
- Second, specifies y_{2t} as the common stochastic trend with innovation e_{2t}

Example: Phillip's triangular representation

$$
\begin{aligned}
T &= 250 & \mathbf{a} &= (1, -1) \\
e_{1t} &= 0.75 e_{1,t-1} + \nu_t & \nu_t &\sim N(0, 0.5^2) \\
e_{2t} &\sim IIDN(0, 0.5^2)
\end{aligned}
$$

```
> library(splus2R)
> T=250
> e= rmvnorm(T, mean=rep(0,2), sd=c(0.5,0.5))
> e1.ar1= arima.sim(n=T,model=list(ar=0.75), innov= e[,1])
> y2= cumsum(e[,2])
> y1= y2 + e1.ar1
```

Example: Phillip's triangular representation



They follow each other closely because their stochastic trend is the same.

Simulating Cointegrated Systems

Trivariate: 1 cointegrated vector, 2 common trends

- Assume $\mathbf{y_t} = (y_{1t}, y_{2t}, y_{3t})'$ is cointegrated,
- and $\mathbf{a} = (1, -a_2, -b_3)'$ is the cointegrating vector

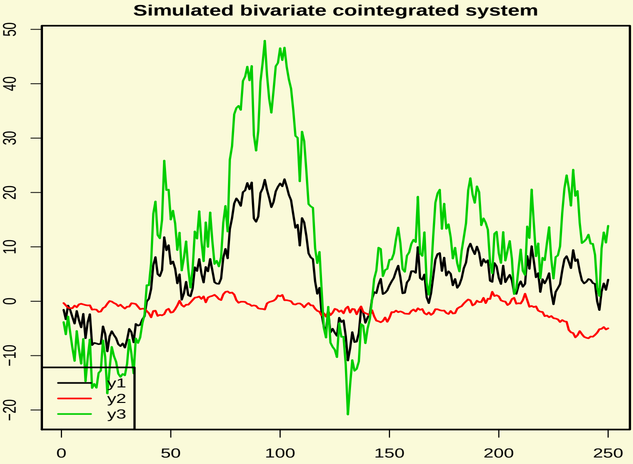
$$
\begin{aligned}
y_{1t} =& a_2 y_{2t} + a_3 y_{3t} + e_{1t} & \qquad e_{1t} \sim I(0) \\
y_{2t} =& y_{2,t-1} + e_{2t} & \qquad e_{2t} \sim I(0) \\
y_{3t} =& y_{3,t-1} + e_{3t} & \qquad e_{3t} \sim I(0)
\end{aligned}
$$

- First eq describes the long-run equilibrium relationship
- Second and third describe the common stochastic trends

Example: Trivariate

$$
\begin{aligned}
T &= 250 & \mathbf{a} &= (1, -0.5, -0.5) \\
e_{1t} &= 0.75 e_{1,t-1} + \nu_t & \nu_t &\sim N(0, 0.5^2) \\
e_{2t} &\sim IIDN(0, 0.5^2) & e_{3t} &\sim IIDN(0, 4)
\end{aligned}
$$

Example: Trivariate



y_2, y_3 are two independent common trends, y_1 is the average of the two trends plus a AR(1) residuals.

Simulating Cointegrated Systems

Trivariate: 2 cointegrated vectors, 1 common stochastic trend

- Assume $\mathbf{y_t} = (y_{1t}, y_{2t}, y_{3t})'$ is cointegrated,
- and $\mathbf{a_1} = (1, 0, -a_{13})'$ and $\mathbf{a_2} = (0, 1, -a_{23})$ cointegrated vectors

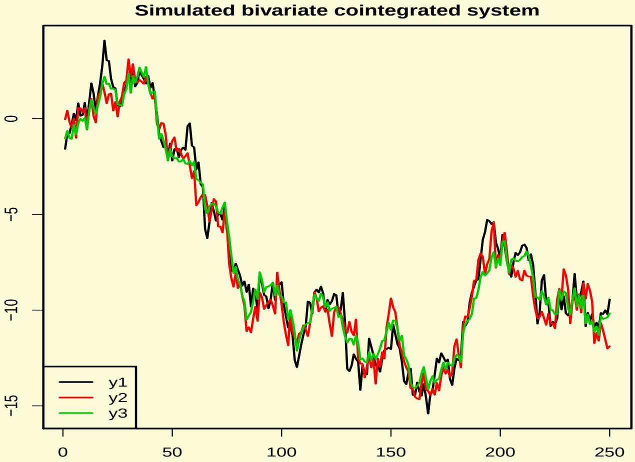
$$
\begin{aligned}
y_{1t} &= a_{13} y_{3t} + e_{1t} & e_{1t} &\sim I(0) \\
y_{2t} &= a_{23} y_{3,t-1} + e_{2t} & e_{2t} &\sim I(0) \\
y_{3t} &= y_{3,t-1} + e_{3t} & e_{3t} &\sim I(0)
\end{aligned}
$$

- First two eqs describe the two long-run equilibrium relationships
- Third describes the common stochastic trend

Example: Trivariate

$$
\begin{aligned}
T &= 250 & \mathbf{a1} &= (1,0,-1) \quad \mathbf{a2} = (0,1,-1) \\
e_{1t} &= 0.75 e_{1,t-1} + \nu_t & \nu_t &\sim N(0, 0.5^2) \\
e_{2t} &= 0.75 e_{2,t-1} + \xi_t & \xi_t &\sim N(0, 0.5^2) \\
e_{3t} &\sim IIDN(0, 0.5^2)
\end{aligned}
$$

Example: Trivariate



Error-correction representation (ECM)

For a bivariate VAR(p),

- Assume $\mathbf{y_t} = (y_{1t}, y_{2t})'$ is cointegrated,
- and $\mathbf{a} = (1, -a_2)'$ is the cointegrating vector

$$\Rightarrow \quad \mathbf{a}'\mathbf{y_t} = y_{1t} - a_2 y_{2t} \quad \text{is I(0)}$$

- Engle and Granger (1987) show that cointegration implies the existence of *error correction model* (ECM) describing the dynamic behaviour of y_{1t} and y_{2t}.

Error-correction representation (ECM)

$$\Delta y_{1t} = \alpha_1 + \pi_1 \underbrace{(y_{1,t-1} - a_2 y_{2,t-1})}_{\mathbf{a}'\mathbf{y}_t}$$

$$+ \underbrace{\sum_j \xi_{11}^j \Delta y_{1,t-j} + \sum_j \xi_{12}^j \Delta y_{2,t-j}}_{\text{short run dynamic adjustment}} + \epsilon_{2t}$$

$$\Delta y_{2t} = \alpha_2 + \pi_2 (y_{1,t-1} - a_2 y_{2,t-1})$$

$$+ \sum_j \xi_{21}^j \Delta y_{1,t-j} + \sum_j \xi_{22}^j \Delta y_{2,t-j} + \epsilon_{2t}$$

SYDDANSK UNIVERSITET

Error Correction Model (ECM)

Intuition: $z_t = y_{t1} - a_2 y_{2t} \sim I(0)$, then $y_{t1} - a_2 y_{2t}$ is the disequilibrium error

- Where the system goes at time $t+1$ depends on the sign and magnitude of the disequilibrium error at time t.
- Short-run dynamics are movements in the short run, modeled in the ECM, that guide the economy towards the long-run equilibrium $y_{1t} = a_2 y_{2t}$

Example: bivariate ECM for stock prices and dividends

- Let S_t be the stock price and D_t the annual dividend
- The dividend yield D_t/S_t, after taking logarithms become $d_t - s_t$
- $\mathbf{y_t} = (s_t, d_t)'$ is I(1).
- However the divident yield $d_t - s_t \sim I(0)$
- $\mathbf{a}' = (1, -1)$
- The theory says that there is a long-run equilibrium in the dividend yield:

$$d_t = s_t + \mu + e_t \qquad e_t \sim I(0)$$

where μ is the mean of the log dividend-price ratio and e_t is an I(0) random variable.

Example: bivariate ECM for stock prices and dividends

The ECM has the form:

$$\Delta s_t = \alpha_s + \pi_s(d_{t-1} - s_{t-1} - \mu) + \epsilon_{st}$$
$$\Delta d_t = \alpha_d + \pi_d(d_{t-1} - s_{t-1} - \mu) + \epsilon_{dt}$$

where $\alpha_s, \alpha_d > 0$.

- The first equation relates the growth rate of stock prices to the lagged disequilibrium error $d_{t-1} - s_{t-1} - \mu$
- The second eq. relates the growth rate of dividends to the lagged of the disequilibrium error
- The reactions to the disequilibrium error are adjusted by π_s, π_d.

Example: bivariate ECM for stock prices and dividends

Let us assume that $\pi_d = 0$ and $\pi_s = 0.5$ (only s_t reacts to the disequilibrium)

Case 1 $d_{t-1} - s_{t-1} - \mu = 0$, then $E(\log s_t | \mathbf{y}_{t-1}) = \alpha_s$ and $E(\log d_t) = \alpha_d$. These quantities represent the growth rates of stock prices and dividends in long-run equilibrium.

Case 2 $d_{t-1} - s_{t-1} - \mu > 0$, then $E(\log s_t | \mathbf{y}_{t-1}) > \alpha_s$. The dividend yield has increased above the long-run mean. The ECM predicts that s_t will grow faster than the equilibrium rate to restore the dividend yield to its long-run mean.

Case 3 $d_{t-1} - s_{t-1} - \mu < 0$. The dividend yield has decreased below its long-run mean. ????

Residual-based tests for cointegration

Let $\mathbf{y}_t$ be a I(1) cointegrated vector: $\mathbf{A}'\mathbf{y_t} \sim I(0)$

Testing for cointegration can be seen as testing for the existence of a long-run equilibria in the elements of $\mathbf{y}_t$. Two scenarios:

1. There is at most one cointegrating vector (Engle and Granger, 1986)
   - Two-step residual-based test
2. There are possibly $0 \leq h < n$ cointegrating vectors (Johansen, 1988)
   - A more complicated procedure to determine the number of cointegrating relationships

Engle and Granger's cointegration test

Once the cointegrating vector is pre-defined, we can test whether the residuals $z_t = \mathbf{a}'\mathbf{y_t}$ are stationary.

1. Test whether each individual element of $\mathbf{y}_t$ is I(1) – DF test or tests of stationarity
2. We construct and scalar $z_t = \mathbf{a}'\mathbf{y_t}$ and we test $H_0 : z_t \sim I(1)$ vs $H_1 : z_t \sim I(0)$.

Conclusion:

If z_t is stationary then $\mathbf{y}_t$ is cointegrated with cointegrating vector $\mathbf{a}'$

Example: PPP

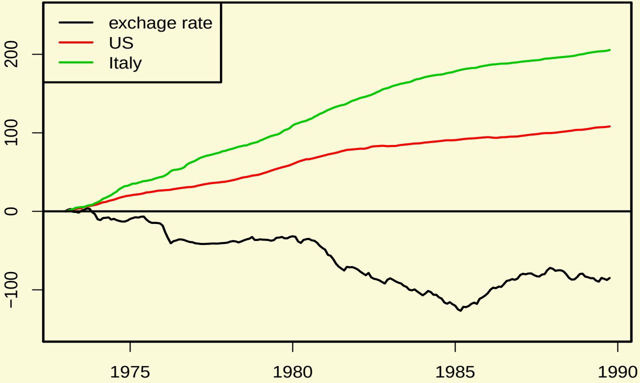
Data from 1973:1 – 1989:10.

- $p_t = 100(\log(P_t) - \log(P_1))$ US consumer price
- p_t^* Italy consumer price
- $s_t = -100(log(S_t) - \log(S_1))$ dollar–lira exchange rate
- $z_t = p_t - s_t - p_t^*$

Example: PPP

```
> load("../data/ppp.rda")
> selection <- window( ppp, start=c(1973,1), end=c(1989,10) )
> ppp.data <- cbind(
+   pstar=100*log(selection[,"PC6IT"]/selection[[1,"PC6IT"]]),
+   p=100*log(selection[,"PZUNEW"]/selection[[1,"PZUNEW"]]),
+   s=-100*log(selection[,"EXRITL"]/selection[[1,"EXRITL"]])
+   )
> ppp.data <- cbind( ppp.data,
+                    z = ppp.data[,"p"] - ppp.data[,"s"] - ppp.data[,"pstar"] )
```

Example: PPP



Example: PPP

Perform DF-test for the three series:

- $E(\Delta p_t) > 0$ so $H_0 : p_t$ unit root process with positive drift vs $H_1 : p_t$ stationary around a deterministic trend

```
> library(urca)
> s.df<-ur.df(as.vector(ppp.data[,"s"]), type="trend", lags=12)
> p.df<-ur.df(as.vector(ppp.data[,"p"]), type="trend", lags=12)
> z.df<-ur.df(as.vector(ppp.data[,"z"]), type="drift", lags=12)
```

Example: PPP

```
> pstar.df<-ur.df(as.vector(ppp.data[,"pstar"]),
+                  type="trend", lags=12)
> attr(pstar.df,"teststat")

               tau3    phi2     phi3
statistic -0.1319633 3.59166 4.249956

> attr(p.df,"teststat")

              tau3     phi2     phi3
statistic -1.954675 2.066691 2.412933

> attr(s.df,"teststat")

              tau3     phi2     phi3
statistic -1.584433 1.219635 1.489674
```

Example: PPP

They are all individually I(1). Theory says that $\mathbf{a}' = (1, -1, -1)$, then

$$z_t = p_t - s_t - p_t^*$$

The trends should be eliminated with this transformation, so we perform the ADF test on z_t with drift and get a statistics -2.04 $>$-2.88 and so the null hypothesis of unit root is accepted.

The series are not cointegrated. At least not with this cointegrating vector

Estimating the cointegrating vector

- If there is no theoretical candidate for $\mathbf{a}$, then it has to be estimated
- We find first a candidate with OLS
- If $z_t = \mathbf{a}'\mathbf{y_t}$ is second moment stationary and ergodic:

$$\frac{1}{T}\sum_t z_t^2 = \frac{1}{T}\sum_t (\mathbf{a}'\mathbf{y_t})^2 \to^p E(z_t^2)$$

- If instead, $\mathbf{a}$ is not the cointegrating vector $\Rightarrow z_t \sim I(1)$ and $\frac{1}{T}\sum_t z_t^2 \to \infty$ as $T \to \infty$
- So we find and estimate of the cointegrating vector by minimising $\frac{1}{T}\sum_t (\mathbf{a}'\mathbf{y_t})^2$ with respect to $\mathbf{a}$

SYDDANSK UNIVERSITET

Estimating the cointegrating vector

If we know for sure that $a_1 \neq 0$, then we set it to $\hat{\mathbf{a}} = (1, -\hat{a}_2, -\hat{a}_3, \ldots, -\hat{a}_n)$ result of the OLS estimation of:

$$y_{1t} = \alpha + a_2 y_{2t} + a_3 y_{3t} + \ldots + a_n y_{tn} + \epsilon_t$$

- These estimates are super consistent (converges at rate T), even if there is endogeneity.
- Their asymptotic distribution is non-normal.
- It might be substantially biased in small samples.
- A better estimator can be found.

SYDDANSK UNIVERSITET

Estimating the cointegrating vector

$$y_{1t} = \alpha + a_2 y_{2t} + a_3 y_{3t} + \ldots + a_n y_{tn} + \epsilon_t$$

- Let us estimate the cointegrating vector with OLS, what occurs if there is no cointegration relation? What does this estimate mean?
- $y_{1t} = \alpha + \mathbf{a_2'} \mathbf{y_{2t}} + \epsilon_t$
- $\hat{\epsilon}_t = y_{1t} - \hat{\alpha} - \hat{\mathbf{a}}_2' \mathbf{y_{2t}}$
- The unit root test on $\hat{\epsilon}_t$ is without drift or trend
- All the elements in the regression are I(1). So the OLS estimates will probably be spurious and $\hat{\epsilon}_t \sim I(1)$
- So the distribution of the test have other type of asymptotic distributions knonw as the Phillips-Ouliaris (PO)
- These distributions also depend on whether $y_{1t}, \mathbf{y}_{2t}$ have drift or not.

Example: PPP

We say that p_t, p_t^*, s_t for the previous example were not cointegrated with cointegrating vector (1, -1,-1). What about for another linear relationship?

```
> ppp.lm<-lm( p ~ 1 + s + pstar, ppp.data )
```

```
> epsilon.hat<-resid(ppp.lm)
> epsilon.hat.df<-ur.df(epsilon.hat, type="none", lags=12)
> attr(epsilon.hat.df,"teststat")
```

```
              tau1
statistic -2.73094
```

$$p_t = 2.71 + 0.05 s_t + 0.53 p_t^* + \hat{\epsilon}_t$$

We are in a case 3 of test for cointegration (Table 19.1). The critical values for the DF test are in Table B.9. There is little evidence of cointegration

Efficient lead/lag estimator

- Stock and Watson (1993) amongst others suggest an asymptotically efficient estimator (equivalent to MLE) for normalised cointegrating vector $\mathbf{a} = (1 - \mathbf{a}_2)$
- Estimate the following augmented the cointegrating regression by OLS

$$y_{1t} = (\alpha + \delta t) + \mathbf{a_2 y_{2t}} + \sum_{j=-p}^{p} \rho_j \Delta \mathbf{y}_{2,t-j} + u_t$$

- The resulted $\hat{\mathbf{a}}_\mathbf{2}$ is called the *dynamic OLS* (DOLS) estimator which is consistent, asymptotically normal and efficient under certain conditions
- Asymptotically valid standard errors are given by the OLS standard errors multiplied by a ratio

Example: Income/consumption