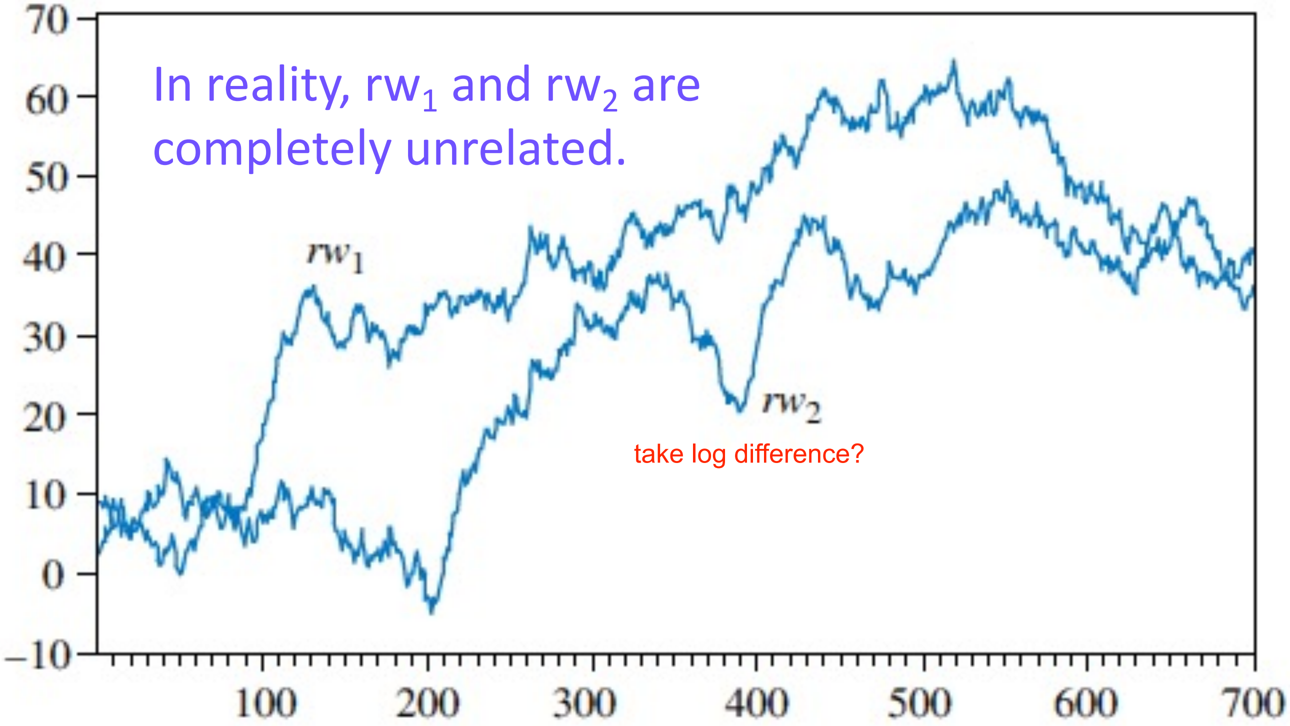
Economics 144
Economic Forecasting

Lecture 15
Cointegration

Dr. Randall R. Rojas

Spurious Regressions



(a) Time series

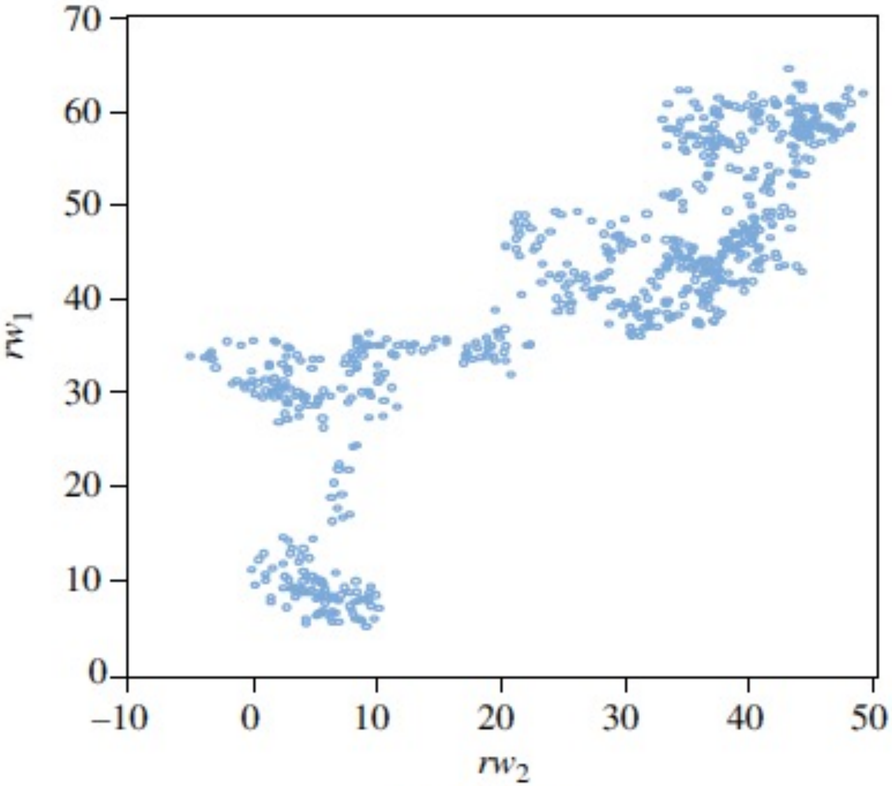
series where it's just random

Random Walk 1 (rw_1): $y_t = y_{t-1} + e_{1t}$

Random Walk 2 (rw_2): $x_t = x_{t-1} + e_{2t}$

$$e_{it} \sim \mathcal{N}(0, 1)$$

Spurious Regressions



(*b*) Scatter plot

- Suppose we did not know that rw_1 and rw_2 were unrelated, so we fit a regression model.

- A simple regression of series one (rw_1) on series two (rw_2) yields:

$$\widehat{rw}_{1t} = 17.818 + 0.842\, rw_{2t}, \qquad R^2 = 0.70$$
$$(t) \qquad\qquad\qquad (40.837)$$

These results are completely meaningless, or spurious.

- The apparent significance of the relationship is false

Spurious Regressions

- When nonstationary time series are used in a regression model, the results may spuriously indicate a significant relationship when there is none
  - In these cases the least squares estimator and least squares predictor do not have their usual properties, and *t*-statistics are not reliable.
  - Since many macroeconomic time series are nonstationary, it is particularly important to take care when estimating regressions with macroeconomic variables.

Unit Root Tests for Stationarity

- There are many tests for determining whether a series is stationary or nonstationary

adf - augmented dickey-fuller test

  - The most popular is the Dickey–Fuller (DF) test

AR1/MA1 - sufficient condition for covar stationary & invertability
- theta=1 - solution is 1 - unit root - would not be stationary

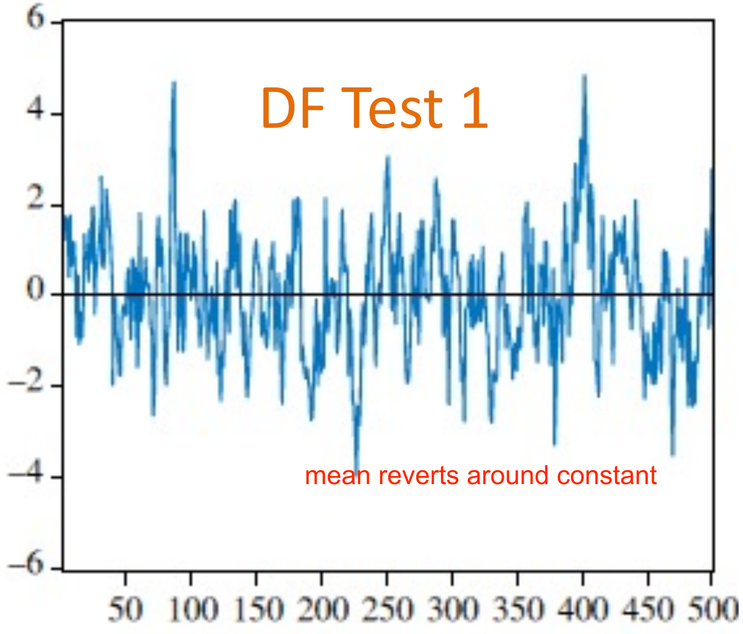
- if data has unit root - problem, if doesn't - stationary

- There are three popular cases:
  - DF Test 1 (No Constant and No Trend): $\Delta y_t = \gamma y_{t-1} + e_t$ ar1 - behave as a slope
  - DF Test 2 (With Constant but No Trend): $\Delta y_t = \alpha + \gamma y_{t-1} + e_t$

more general case

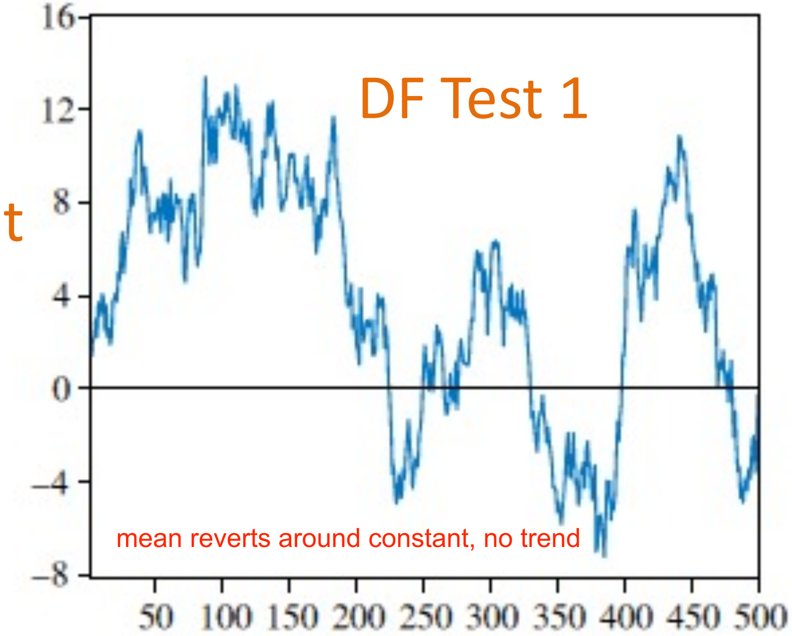
  - DF Test 3 (With Constant and With Trend): $\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + e_t$

has some growth

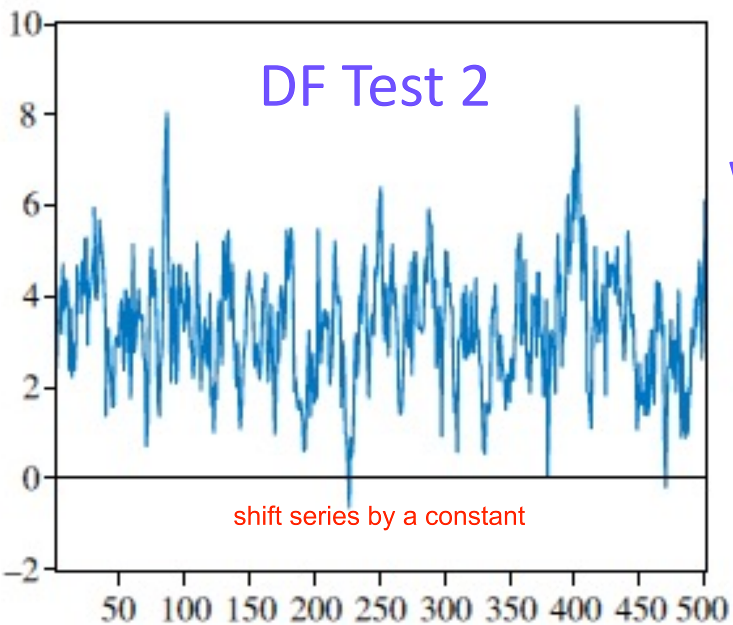


(a) $y_t = 0.7y_{t-1} + v_t$

No Constant
and
No Trend

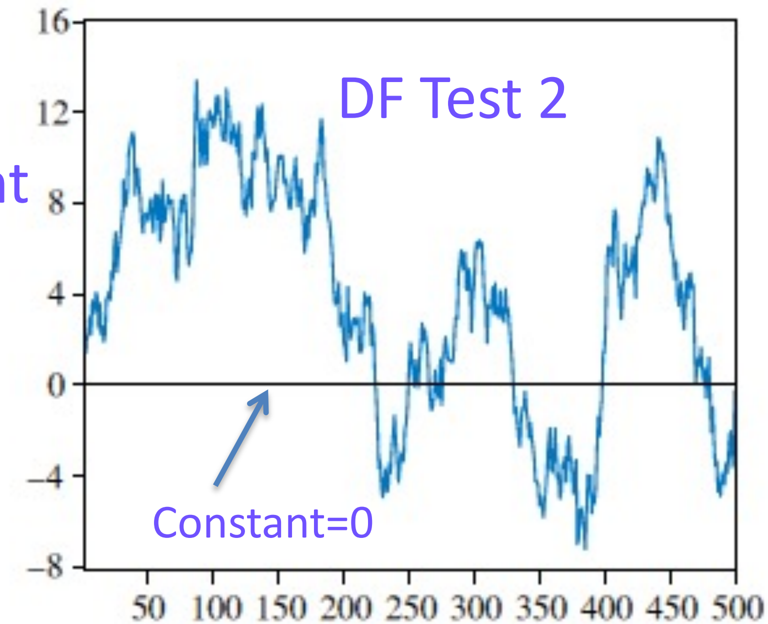


(d) $y_t = y_{t-1} + v_t$

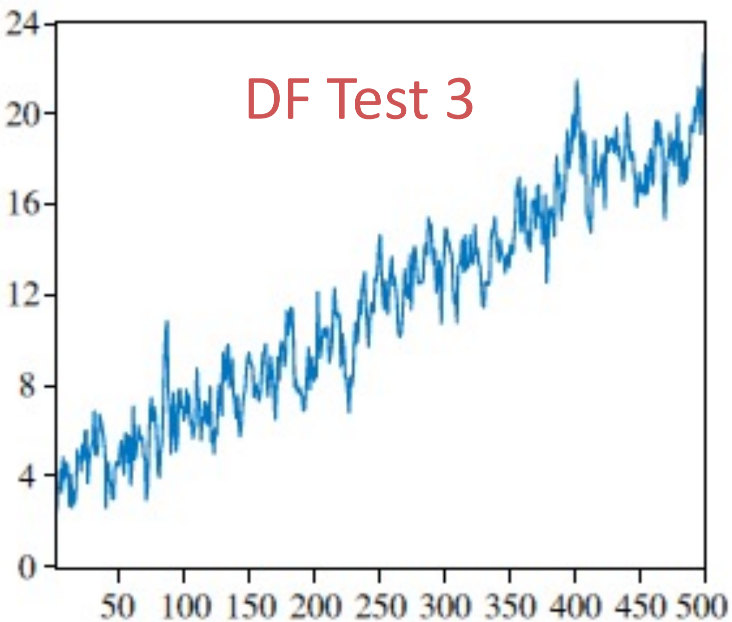


(b) $y_t = 1 + 0.7y_{t-1} + v_t$

With Constant
but
No Trend



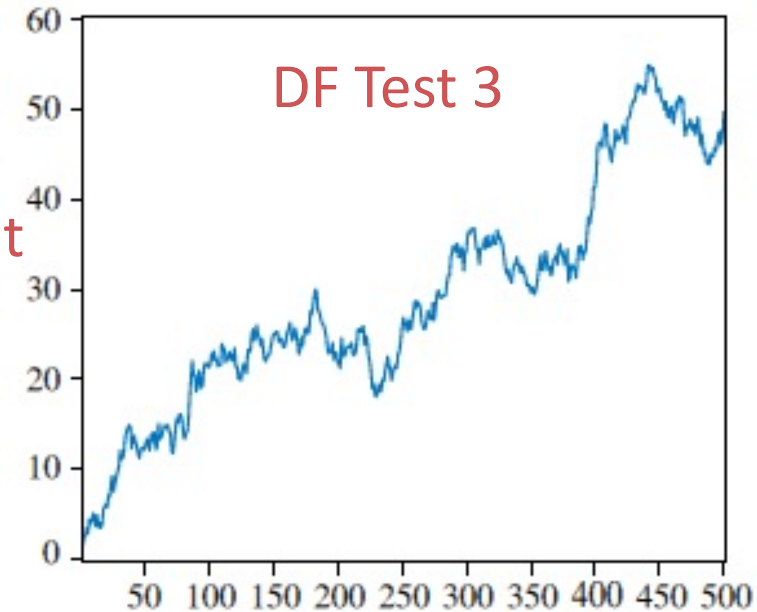
(d) $y_t = y_{t-1} + v_t$



(c) $y_t = 1 + 0.01t + 0.7y_{t-1} + v_t$

With Constant
and
With Trend

greater concern for us



(e) $y_t = 0.1 + y_{t-1} + v_t$

Unit Root Tests for Stationarity

- To test the hypothesis in all three cases, we simply estimate the test equation by least squares and examine the t-statistic for the hypothesis that $\gamma = 0$. null under the specific model
  - Unfortunately this t-statistic no longer has the t-distribution.

perform directly and evaluate if it's stationary or not

  - Instead, we use the statistic often called a τ (tau) statistic.

all test for stationarity

Model	1%	5%	10%
$\Delta y_t = \gamma y_{t-1} + v_t$	−2.56	−1.94	−1.62
$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$	−3.43	−2.86	−2.57
$\Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + v_t$	−3.96	−3.41	−3.13
Standard critical values	−2.33	−1.65	−1.28

Note: These critical values are taken from R. Davidson and J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press, p. 708.

Unit Root Tests for Stationarity

- To carry out a one-tail test of significance, if τ_c is the critical value obtained from the DF Table, we reject the null hypothesis of nonstationarity if $\tau \leq \tau_c$
  - If $\tau > \tau_c$ then we do not reject the null hypothesis that the series is nonstationary
- An important extension of the Dickey–Fuller test allows for the possibility that the error term is autocorrelated. These tests are referred to as augmented Dickey–Fuller (ADF) tests. encompasses broader range of possibility - done internally
  - When $\gamma = 0$, in addition to saying that the series is nonstationary, we also say the series has a unit root.
  - In practice, we always use the augmented Dickey–Fuller test.

The Dickey-Fuller Testing Procedure

- Plot the time series of the variable and select a suitable Dickey-Fuller test based on a visual inspection of the plot

  avoid it - feed data to df test in r - figure out which one is appropriate and carry out which test to run

  - If the series appears to be wandering or fluctuating around a sample average of zero, use DF Test 1.
  - If the series appears to be wandering or fluctuating around a sample average which is nonzero, use DF Test 2.
  - If the series appears to be wandering or fluctuating around a linear trend, use DF Test 3.

Cointegration

spurious regression & root test

- When we can find a linear combination of integrated variables that is stationary, we say that these variables are cointegrated.

- If y_t and x_t are cointegrated, we expect them to share similar stochastic trends.

- Cointegration is the statistical notion that corresponds to the economic notion of equilibrium.

Cointegration

first difference

- For example, assume y_t and x_t are nonstationary $\mathrm{I}(1)$ variables, then $e_t = y_t - \beta_1 - \beta_2 x_t$ is a stationary $\mathrm{I}(0)$ process.

no longer need to take difference

- We can test whether y_t and x_t are cointegrated by using a Dickey-Fuller test (for stationarity) on the LS residuals (since we cannot observe e_t).

- Basically, if the residuals are stationary, then y_t and x_t are cointegrated.

Cointegration

- Def: Cointegration Relation $y_t = \beta_1 - \beta_2 x_t + e_t$
- Def: Disequilibrium Error e_t measures how far the system y_t, x_t is from the equilibrium path. residual from the regression
- Equilibrium: $e_t = 0$
- Disequilibrium: $e_t > 0$ or $e_t < 0$
- For Stationary data we can use models that capture short-term features. e.g. moving average
- For Non-Stationary data we can use models that capture long-term features. e.g. ar
- Q: How can we integrate short- and long-term dynamics in a time series? combine models that capture short term+long term dynamics
- A: The Error Correction Model (based on cointegration)

combine short term fluctuation with long term

Cointegration

- The test for stationarity of the residuals is based on the test equation:

asking a lot

$$\Delta\hat{e}_t = \gamma\hat{e}_{t-1} + v_t$$

model it as ar1

errors should be stationary by themselves

  - The regression has no constant term because the mean of the regression residuals is zero.
  - We are basing this test upon estimated values of the residuals.
  - The rhs of the equation can include higher order lags if needed to eliminate any dynamics in v_t.

Cointegration

- There are three sets of critical values
  - Which set we use depends on whether the residuals are derived from:

$$Equation\,1:\quad \hat{e}_t = y_t - bx_t$$

dicky fuller test 1-2-3

$$Equation\,2:\quad \hat{e}_t = y_t - b_2 x_t - b_1$$

error - vector model - combine short term and long term variables

$$Equation\,3:\quad \underline{\hat{e}_t = y_t - b_2 x_t - b_1 - \hat{\delta} t}$$

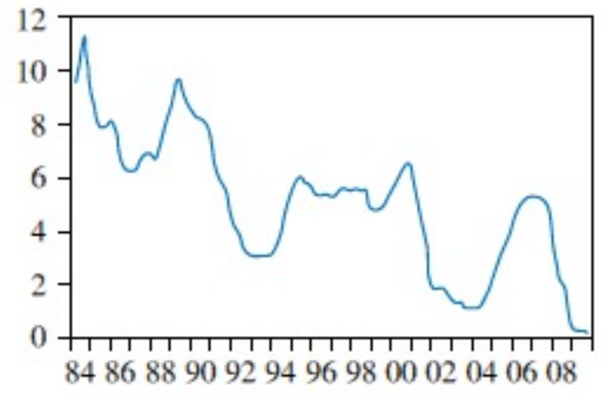
Regression model	1%	5%	10%
(1) $y_t = \beta x_t + e_t$	−3.39	−2.76	−2.45
(2) $y_t = \beta_1 + \beta_2 x_t + e_t$	−3.96	−3.37	−3.07
(3) $y_t = \beta_1 + \delta t + \beta_2 x_t + e_t$	−3.98	−3.42	−3.13

Note: These critical values are taken from J. Hamilton (1994), *Time Series Analysis*, Princeton University Press, p. 766.

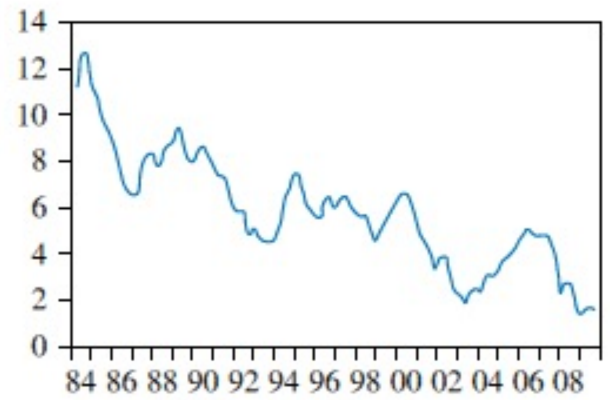
Cointegration

- As an example, consider the two interest rate series (both are non stationary I(1)):
  - The federal funds rate (Ft)

first difference should be covariance stationary



  - The three-year bond rate (Bt)



Cointegration

- Consider the estimated model:

$$\hat{B}_t = 1.140 + 0.914F_t, \quad R^2 = 0.881$$
$$(t) \quad (6.548) \quad (29.421)$$

- The unit root test for stationarity in the estimated residuals is:

take first diff of residual, regressed on first lag and first lag of difference
- verify that we have reduced this to a cointegrated relatiosnhip

$$\Delta\hat{e}_t = -0.225\hat{e}_{t-1} + 0.254\Delta\hat{e}_{t-1}$$
$$(tau) \quad (-4.196)$$

Mathematically solving for that residual

- Use Equation 2 since: $\widehat{e}_t = B_t - 1.140 - 0.914F_t$

Cointegration

- The null and alternative hypotheses in the test for cointegration are:

$$H_0: \text{the series are not cointegrated} \Leftrightarrow \text{residuals are nonstationary}$$

$$H_1: \text{the series are cointegrated} \Leftrightarrow \text{residuals are stationary}$$

- Similar to the one-tail unit root tests, we reject the null hypothesis of no cointegration if $\tau \leq \tau c$, and we do not reject the null hypothesis that the series are not cointegrated if $\tau > \tau c$.

Result: Reject $H_0 \rightarrow$ Federal Funds and Bond Rates are cointegrated!

Implication: It means that when the Fed implements monetary policy by changing the federal funds rate, the bond rate will also change thereby ensuring that the effects of monetary policy are transmitted to the rest of the economy.

With a cointegrated relationship, we're gonna use it for forecast

Error Correction Equation

combine short term and long term relationship

- Consider a general model that contains lags of y_t and x_t.

- Namely, the autoregressive distributed lag model (ARDL), except the variables are nonstationary:

$$y_t = \delta + \theta_1 y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + v_t$$

where v_t are the residuals.

Error Correction Equation

- If y_t and x_t are cointegrated, it means that there is a long-run relationship between them
  - To derive this exact relationship, we set

    $y_t = y_{t-1} = y,\ x_t = x_{t-1} = x$ and $v_t = 0$
  - Imposing this concept in the ARDL, we obtain:

    $\underbrace{y = \beta_1 + \beta_2 x,}_{\text{Cointegrating (or long-run) relationship between } x \text{ and } y}$ where $\beta_1 = \delta/(1-\theta_1)$ and

    $\beta_2 = (\delta_0+\delta_1)/(1-\theta_1)$

long run equilibrium equation

Error Correction Equation

- Add the term $-y_{t-1}$ to both sides of the equation:

$$y_t - y_{t-1} = \delta + \left(\theta_1 - 1\right) y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + v_t$$

- Add the term $-\delta_0 x_{t-1} + \delta_0 x_{t-1}$:

$$\Delta y_t = \delta + \left(\theta_1 - 1\right) y_{t-1} + \delta_0 \left(x_t - x_{t-1}\right) + \left(\delta_0 + \delta_1\right) x_{t-1} + v_t$$

  – Manipulating this we get:

$$\Delta y_t = \left(\theta_1 - 1\right)\left(\frac{\delta}{\left(\theta_1 - 1\right)} + y_{t-1} + \frac{\left(\delta_0 + \delta_1\right)}{\left(\theta_1 - 1\right)} x_{t-1}\right) + \delta_0 \Delta x_t + v_t$$

$$\boxed{\Delta y_t = -\alpha\left(y_{t-1} - \beta_1 - \beta_2 x_{t-1}\right) + \delta_0 \Delta x_t + v_t}$$

short term dynamics on the overall long term dynamics

Error Correction Equation

- The expression $\Delta y_t = -\alpha\left(y_{t-1} - \beta_1 - \beta_2 x_{t-1}\right) + \delta_0 \Delta x_t + v_t$ is called an error correction equation

lag 1 of y and x

- This is a very popular model because:
  - It allows for an underlying or fundamental link between variables (the long-run relationship).
  - It allows for short-run adjustments (i.e. changes) between variables, including adjustments to achieve the cointegrating relationship.
  - It also shows we can work with I(1) variables (y_{t-1}, x_{t-1}) and I(0) variables ($\Delta y_{t-1}, \Delta x_{t-1}$) in the same equation .

Error Correction Equation

- For the bond and federal funds rates example, we have:

one way to obtain co-integrated relationship

$$\Delta\hat{B}_t = -0.142\left(B_{t-1} - 1.429 - 0.777F_{t-1}\right) + 0.842\Delta F_t - 0.327\Delta F_{t-1}$$

$$(t) \qquad (2.857) \qquad\qquad\qquad\qquad (9.387) \qquad (3.855)$$

  – The estimated residuals are

$$\underline{\hat{e}_{t-1} = \left(B_{t-1} - 1.429 - 0.777F_{t-1}\right)}$$

Error Correction Equation

- The result from applying the ADF test for stationarity is:

$$\Delta\hat{e}_t = -0.169\hat{e}_{t-1} + 0.180\Delta\hat{e}_{t-1}$$
$$(t) \quad (-3.929)$$

- Comparing the calculated value (-3.929) with the critical value, we reject the null hypothesis and conclude that (B, F) are cointegrated.

For Next Class

- Readings about today's class:
  Chapter 12[a]

- Review Exercises / Problems:
  Chapter 13[a]: 1, 3, 5, 6

- Readings for next class:
  Special Topics: ARCH/GARCH