# Nonstationarity and cointegration
# (SW Sections 14.6,14.7 and 16.4)

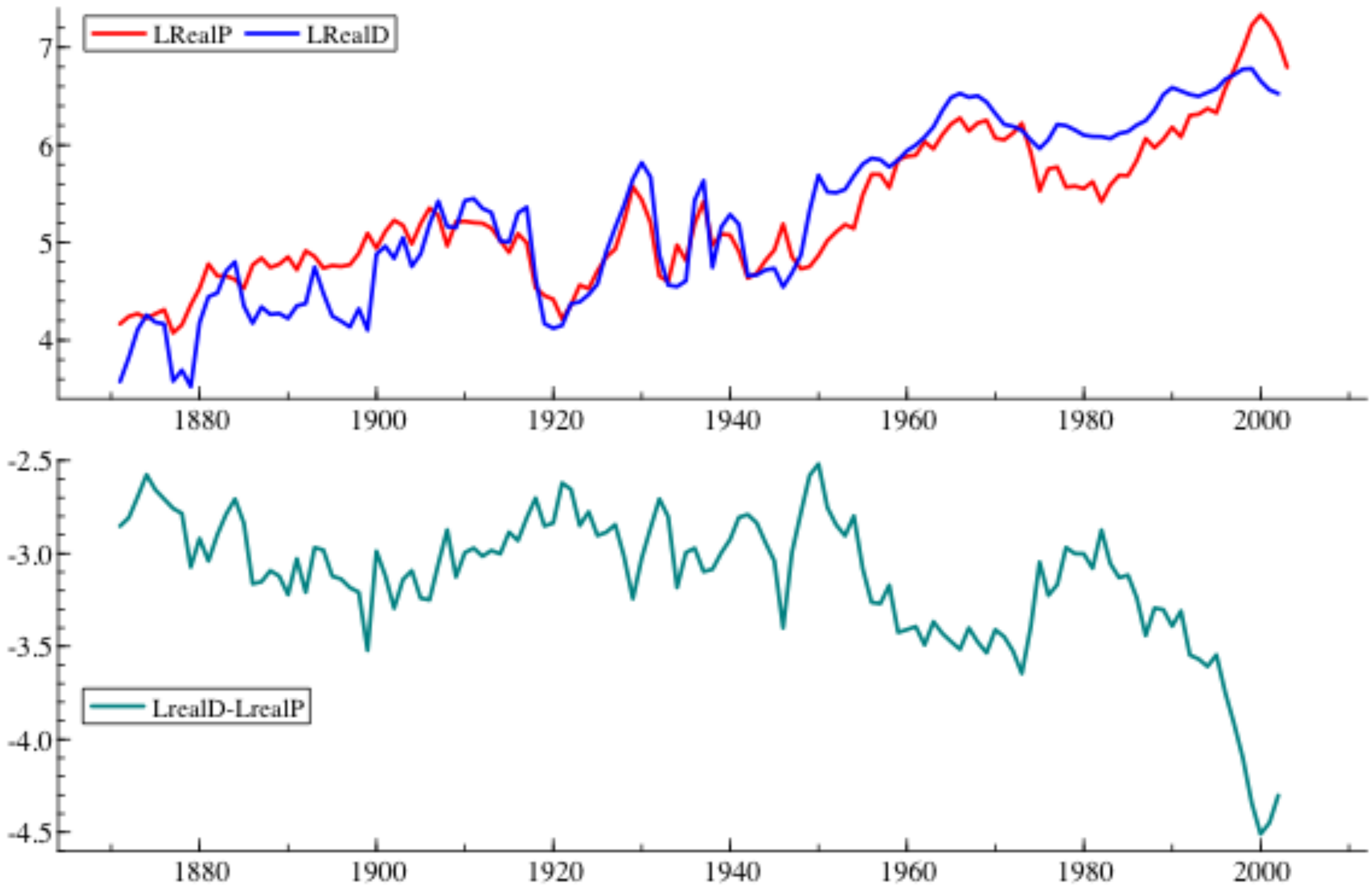
***Economic time series*** are often not well-described by the assumption of stationarity

- GDP, stock prices, consumer prices exhibit trends
- Inflation, interest rates, exchange rates exhibit multi-period swings due to different policy regimes

But series could have ***common trends***

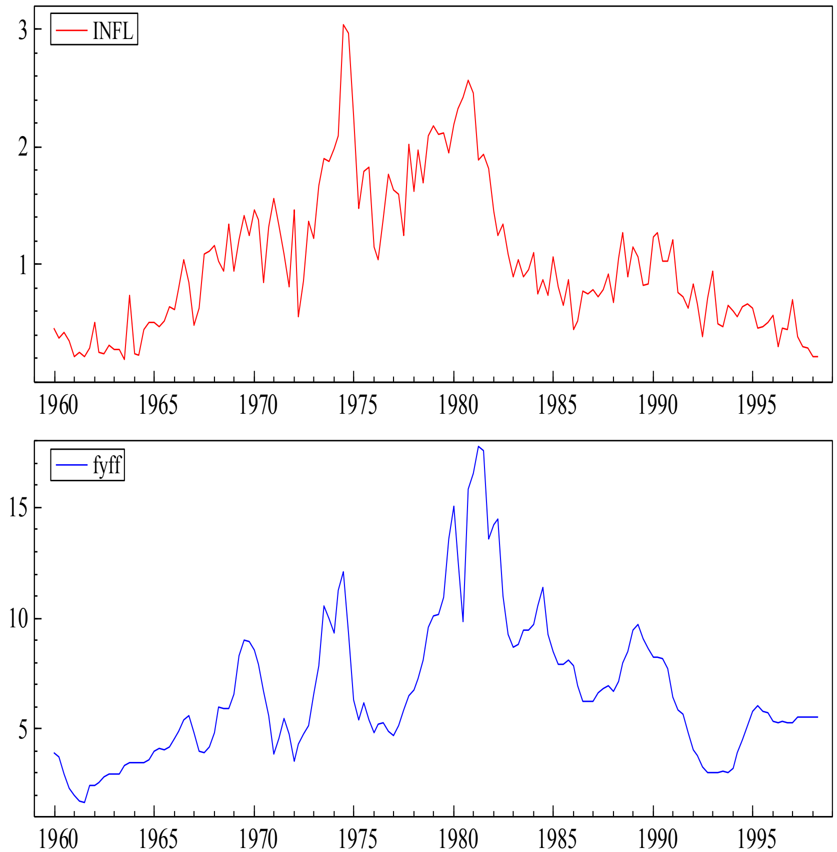
- Stock prices and dividends
- Long versus short maturity bond rates

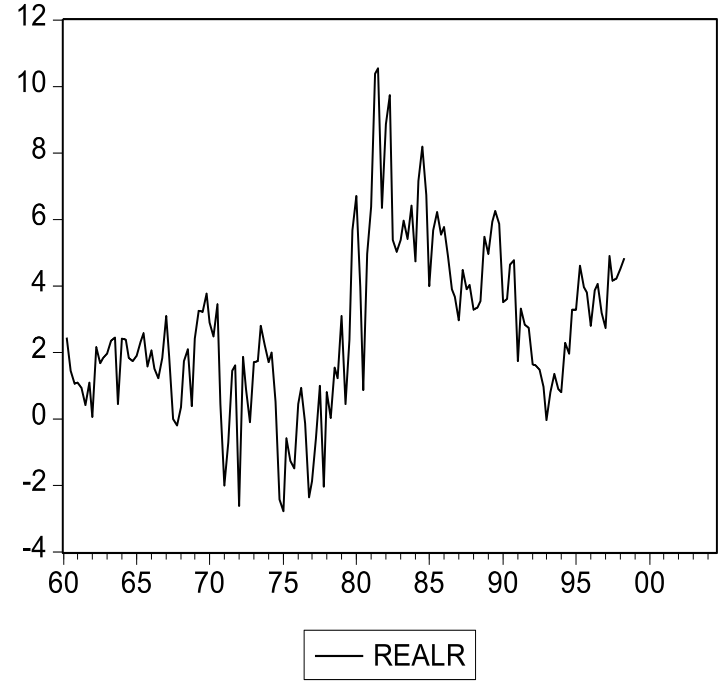
# Example #1 Real US stock prices and dividends (S&P)




Annual log real US Stock Prices (red), Dividends (blue) and log Dividend-Price ratio (green). Data from Figure 1 in R. Shiller "From Efficient Markets Theory to Behavioral Finance," J. Econ. Pers. 2003

# Example #2: Regime shifts in US monetary policy?




US inflation (red), Federal Funds rate (blue) and real interest rate (black)

- Is there a regime shift in US monetary policy in 1980?

# Nonstationarity I: Trends
# (SW Section 14.6)

Overview

- Stochastic versus deterministic trends
- The Random Walk model and unit roots
- *Difference* versus *trend stationarity*
- Problems caused by stochastic trends
- Testing for unit roots in AR models:
  The Augmented Dickey-Fuller (ADF) test

# What is a trend

- A trend is a persistent long-term movement of a variable over time
- It could be deterministic or stochastic
- A ***deterministic trend*** is a deterministic function of time
  - for example $\alpha_1 * t$, where $\alpha_1$ is some constant
- If a time series exhibits stationary deviations from a deterministic trend, it is called ***trend stationary***
  - e.g., AR(p) + DT:

$$Y_t = \beta_0 + \alpha_1 t + \sum_{i=1}^{p} \beta_i Y_{t-i} + u_t$$

## Stochastic trend – Random Walk model

- The (Pure) Random Walk model is defined as

$$Y_t = Y_{t-1} + u_t, \qquad t = 1,2,...$$

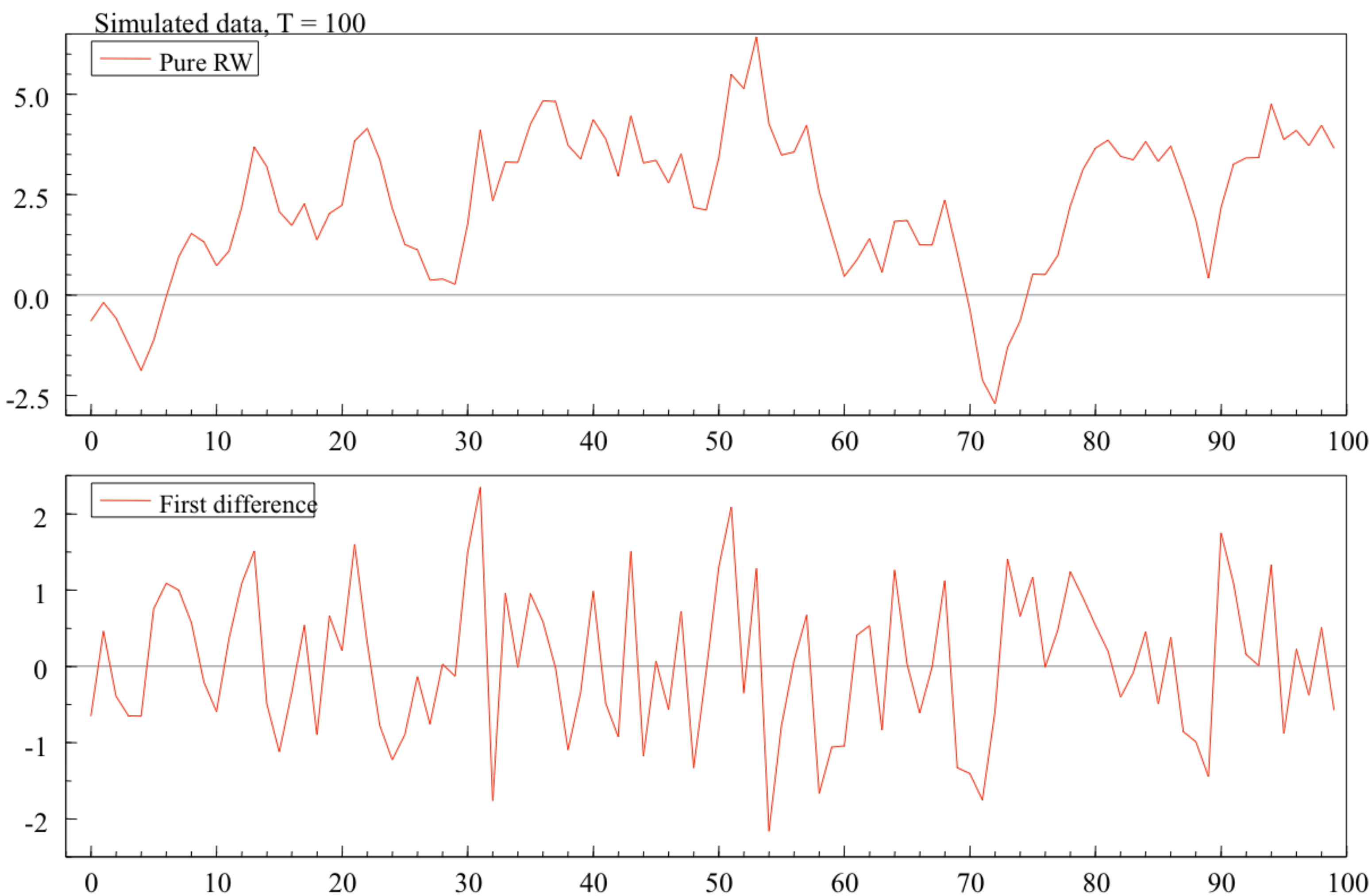
where $u_t$ is serially uncorrelated with mean 0

- This implies

$$Y_t = Y_0 + \sum_{s=1}^{t} u_s, \qquad t = 1,2,...$$

- So, $Y_t$ grows by *random* increments $\Delta Y_t = u_t$
- Contrast this with a deterministic trend that has *nonrandom* increments: $\Delta(\alpha_1\, t) = \alpha_1$
- $Y_t$ is also called *integrated of order* 1, denoted **I(1)**

# Graphical illustration using simulated data




- RW has no tendency to revert to the mean, can wander upwards or downwards for ever

## Random Walk with a drift

- A time series can exhibit both det. and stoch. trends
- A simple model of this is a RW with a drift:

$$Y_t = \alpha_1 + Y_{t-1} + u_t, \qquad t = 1,2,\ldots$$

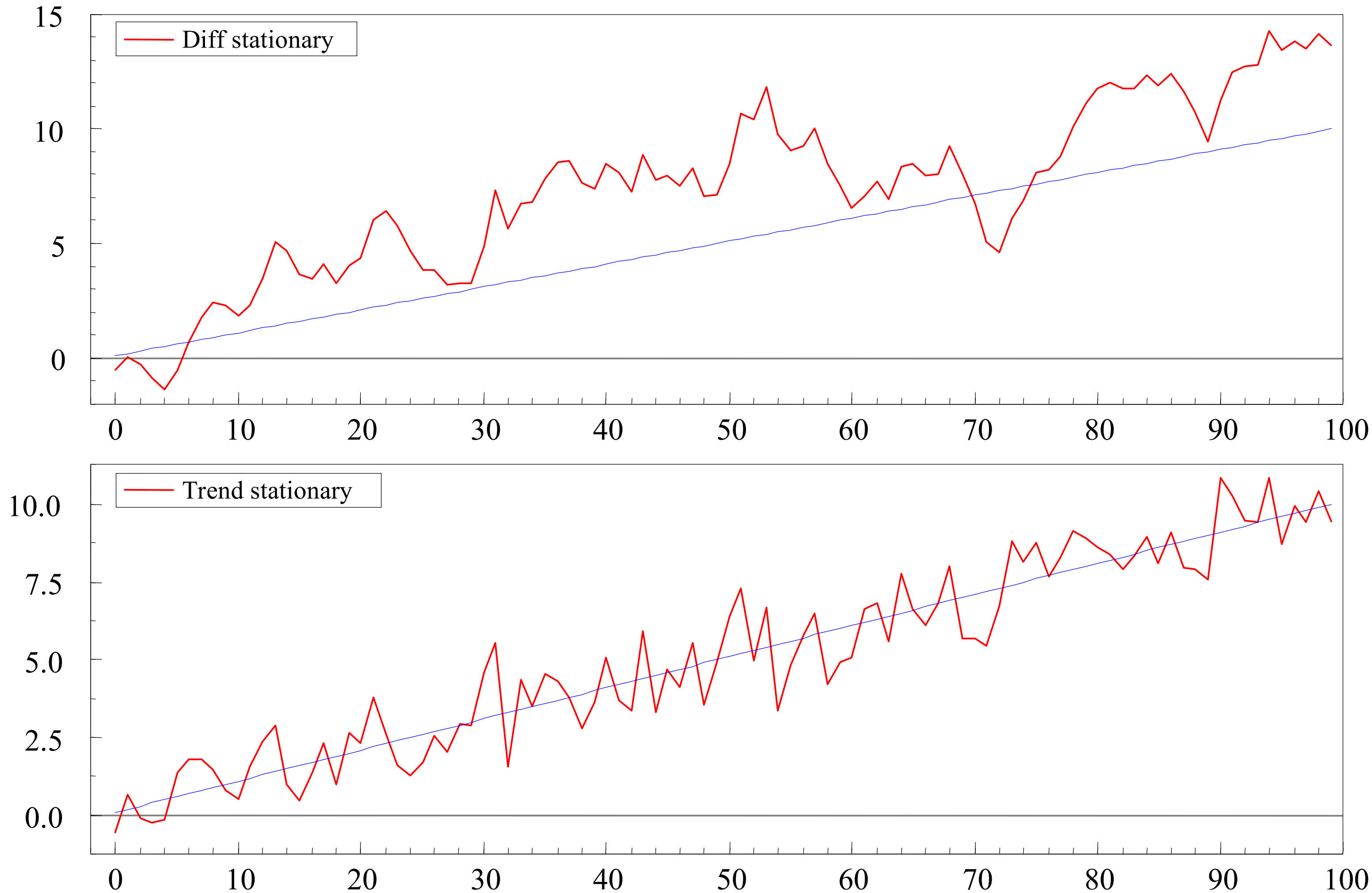
- Assuming $Y_0 = 0$ for simplicity and solving backwards

$$Y_t = \underbrace{\alpha_1 t}_{\text{DT}} + \underbrace{\sum_{s=1}^{t} u_s}_{\text{ST}}, \qquad t = 1,2,\ldots$$

$$E\left(Y_t\right) = \alpha_1 t,\ Var\left(Y_t\right) = \sigma^2 t,\ \operatorname{cov}\left(y_t, y_{t-s}\right) = \sigma^2\left(t-s\right)$$

- Hence, RW model is *nonstationary*

# Trend stationary versus different stationary series




- Simulated using same random errors $u_t$
- RW exhibits persistent deviations from DT line

# Detrending

- It is easy to remove a deterministic trend
- Regress $Y_t$ on deterministic function of time and take the residuals

$$\tilde{Y}_t = Y_t - \hat{\alpha}_{0,OLS} - \hat{\alpha}_{1,OLS} t$$

- Only way to remove stochastic trend is by *differencing*
  - Unless there is *cointegration*, more on this later…
- The terminology "$Y_t$ is integrated of order $q$" means $Y_t$ needs to be differenced $q$ times to remove the stochastic trend
  - $Y_t$ is I(1) means $\Delta Y_t$ is I(0)
  - $Y_t$ is I(2) means $\Delta^2 Y_t = \Delta(\Delta Y_t)$ is I(0)

# Problems caused by stochastic trends

- Autoregressive coefficients biased towards zero
  - RW model is special case of AR(1) with $\beta_1 = 1$
  - This is also called a "unit root" in time series jargon
  - We know that when sample is large, and $\beta_1 = 1$

$$E\left(\hat{\beta}_1^{OLS}\right) \approx 1 - \frac{5.3}{T}$$

- Distribution of OLS estimator and t-statistic is *not* Normal *even in large samples*!
- So, cannot use standard normal critical values and confidence intervals

## Spurious regression

- Stochastic trends can make two unrelated time series to appear related

$$\widehat{U.S.\ Inflation}_t = -37.78 - 3.83\ \ln(\textit{Japanese}\ \mathrm{GDP}_{t-1}),$$

$$(3.99)\quad (0.36), \qquad \bar{R}^2 = 0.56$$

estimated over 1965-1981

$$\widehat{U.S.\ Inflation}_t = 31.20 - 2.17\ \ln(\textit{Japanese}\ \mathrm{GDP}_{t-1}),$$

$$(10.41)\quad (0.80), \qquad \bar{R}^2 = 0.08$$

estimated over 1982-2004

## "Predicting" the stock market

- Regressing real log US stock prices on Var1 and Var2:

```
The estimation sample is: 1871 to 2003
                   Coefficient  Std.Error  t-value  t-prob Part.R^2
 Constant              4.26182     0.1125     37.9   0.000   0.9163
 Var1                -0.153206    0.01503    -10.2   0.000   0.4423

 sigma                0.545949  RSS                 39.0459599
 R^2                  0.442337  F(1,131) =    103.9 [0.000]**

                   Coefficient  Std.Error  t-value  t-prob Part.R^2
 Constant              4.47479     0.1720     26.0   0.000   0.8378
 Var2                0.0985147    0.01929     5.11   0.000   0.1661

 sigma                0.667612  RSS                 58.3874525
 R^2                  0.166098  F(1,131) =    26.09 [0.000]**
```

- Var1 and Var2 are pure random walks generated by my computer!

## Consequences of spurious regression

- Stochastically trending processes will tend to correlate with any other process that exhibits a trend
- The OLS estimator does not converge to any given point – it is random even in large samples
- t and F statistics diverge, so we will be spuriously rejecting the null of no relationship as sample increases
- The $R^2$ is also random in large samples
- But not all regressions involving stochastically trending variables are spurious: the exception is **cointegration**

## Detecting Stochastic trends: Testing for a unit root

- Subtract $Y_{t\text{-}1}$ from both sides of an AR(1) model to get

$$\Delta Y_t = \beta_0 + \delta\, Y_{t-1} + u_t, \;\; \delta = \beta_1 - 1$$

- Test $H_0 : \delta = 0$ against $H_1 : \delta < 0$
  - Note: one-tailed test
  - Null: stochastic trend (unit root)
  - Alternative: $Y_t$ is stationary
- Use the usual t-statistic, but ***not*** Normal critical values
- We have to use different critical values from "Dickey-Fuller" distribution table (see below)
- The resulting test is called ***Dickey-Fuller*** test

# Allowing for deterministic trends under alternative

- Standard Dickey-Fuller model does not allow any deterministic trends under the alternative hypothesis
- If $Y_t$ is trend-stationary, test will be biased in favour of a unit root
  - only way the model can fit the trend is with a unit root
- To avoid that, we use the DF model **with a trend**

$$\Delta Y_t = \beta_0 + \gamma\, t + \delta\, Y_{t-1} + u_t$$

- Same $H_0$ and $H_1$ as before, but $H_1$ now means "$Y_t$ is trend-stationary", not "$Y_t$ is stationary"
- The price is we need **larger** (in abs value) **critical values**

# Augmented Dickey Fuller test

- The DF critical values are correct (in large samples) only if the error in the DF regression is *serially uncorrelated*
- To ensure this we "augment" the DF regression model by including lags of $\Delta Y_t$ on the RHS:

$$\Delta Y_t = \beta_0 + (\gamma\ t) + \delta Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + u_t$$

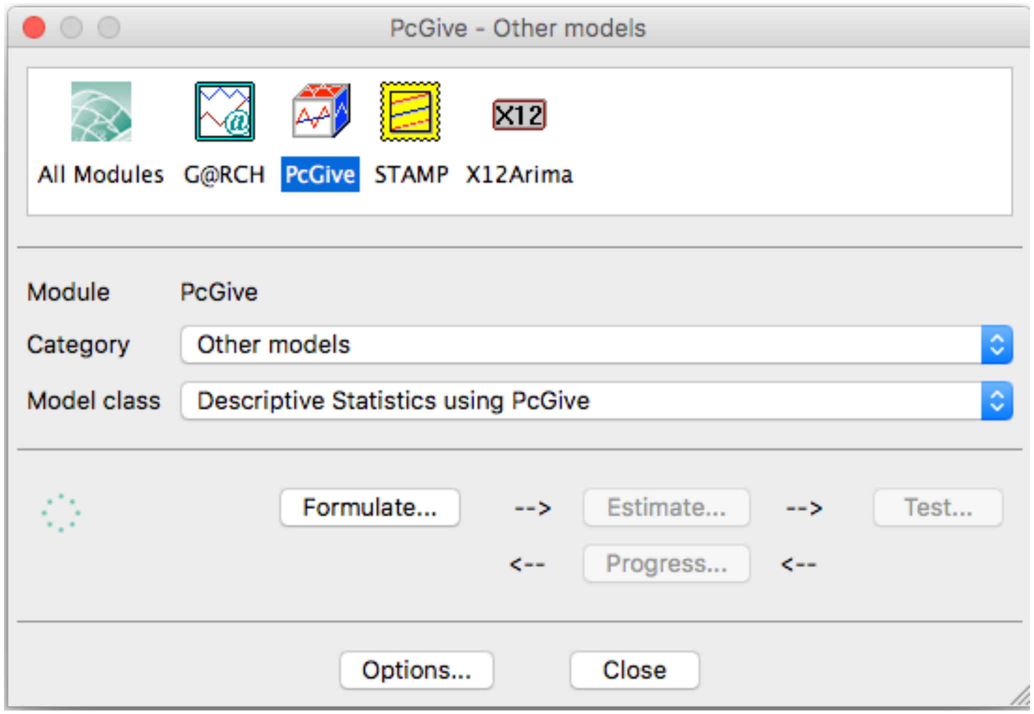
- $H_0$, $H_1$ and critical values are the same as before, but model is called ADF regression and test is ADF test
- Need to use enough lags ($p$ sufficiently large) for the residuals $u_t$ to be serially uncorrelated

# Large sample critical values of ADF test

| | 10% | 5% | 1% |
|---|---|---|---|
| Constant Only | -2.57 | -2.86 | -3.43 |
| Constant + time trend | -3.12 | -3.41 | -3.96 |
| Normal Distribution | -1.28 | -1.645 | -2.33 |

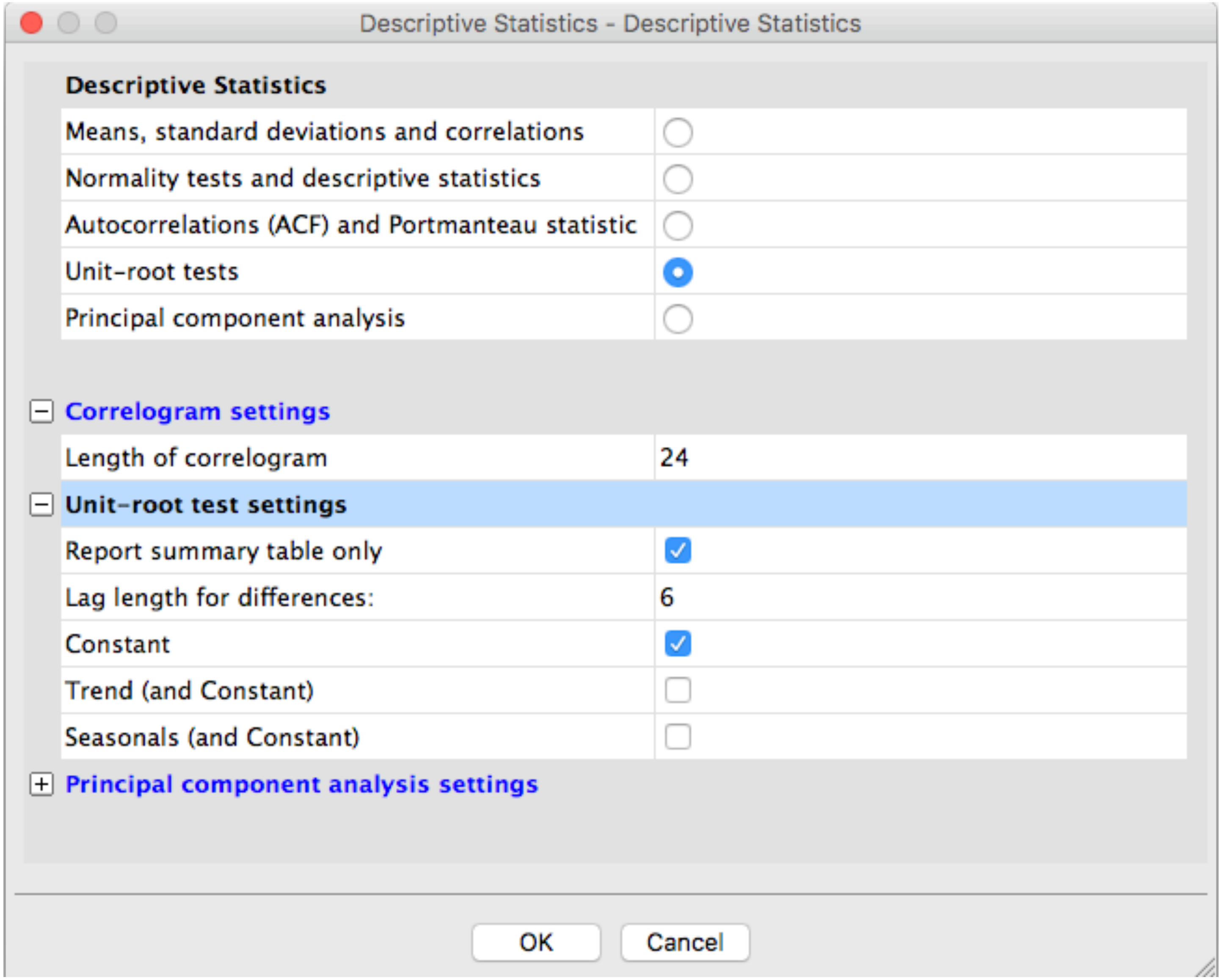
# *Example: Does US inflation have a unit root* – OxMetrics

- Upload the data file StockWatsonUSA.xlsx
- Go to “Model\Model”, choose “Category: Other models” and “Model class: Descriptive Statistics Using PcGive”




- Hit “Formulate”

- Next, select “Inf”, hit “OK”, and then select “Unit root tests”, and select type (Constant) and lag length (6)



- Hit “OK”, select sample 1960:I to 2016:IV, and “OK”

```
Unit-root tests
The dataset is: /Users/sophocles/Dropbox/Teaching/Oxford/Department/QE/2017/!
The sample is: 1960(1) - 2016(4) (235 observations and 1 variables)

Inf: ADF tests (T=228, Constant; 5%=-2.87 1%=-3.46)
D-lag    t-adf      beta Y_1    sigma   t-DY_lag  t-prob       AIC  F-prob
  6     -2.423       0.92466   0.9994      1.125  0.2619   0.03332
  5     -2.280       0.92985    1.000   -0.06604  0.9474   0.03028  0.2619
  4     -2.324       0.92953   0.9978      2.406  0.0169   0.02153  0.5311
  3     -2.001       0.93923    1.008     -1.493  0.1368   0.03850  0.0737
  2     -2.235       0.93265    1.011     -2.697  0.0075   0.03968  0.0572
  1     -2.697       0.91881    1.025     -4.106  0.0001   0.06287  0.0061
  0     -3.638**     0.88986    1.061                       0.1263  0.0000
```

- Seven models have been estimated with $p$=0,…,6
- For efficiency, select most parsimonious model:
    - minAIC yields $p$=4; F-test gives $p$=4 at 10% level
- In all cases, we **accept** the null of a unit root

- Does this mean that inflation is integrated of order 1?
- For this to be the case, ΔInflation must be I(0), i.e., must not have a unit root
- Repeat ADF tests for DInf…

DInf: ADF tests (T=228, Constant; 5%=-2.87 1%=-3.46)

| D-lag | t-adf | beta Y_1 | sigma | t-DY_lag | t-prob | AIC | F-prob |
|---|---|---|---|---|---|---|---|
| 6 | -6.339** | -0.62495 | 1.007 | 1.557 | 0.1210 | 0.04871 | |
| 5 | -6.193** | -0.47339 | 1.010 | -0.7654 | 0.4449 | 0.05089 | 0.1210 |
| 4 | -7.272** | -0.55325 | 1.009 | 0.4244 | 0.6717 | 0.04477 | 0.2239 |
| 3 | -8.048** | -0.51021 | 1.008 | -2.095 | 0.0373 | 0.03680 | 0.3649 |
| 2 | -11.86** | -0.75495 | 1.015 | 1.791 | 0.0747 | 0.04752 | 0.1120 |
| 1 | -14.88** | -0.56805 | 1.020 | 3.097 | 0.0022 | 0.05296 | 0.0586 |
| 0 | -20.58** | -0.30475 | 1.039 | | | 0.08592 | 0.0028 |

- We reject $H_0$ that DInf has a unit root at 1% level
- So we conclude that inflation is I(1)

# Some caution about unit root tests

- ADF test has null hypothesis that series is nonstationary
- Accepting hypothesis of unit root can be due to Type II error
  - ADF test have notoriously little power to distinguish between unit roots and very persistent but stationary alternatives
- SW chapter 16 discusses more powerful tests (beyond scope of this course)
- Economic theory may often provide guidance, e.g., PPP, UIP

# Nonstationarity II: Breaks
# (SW Section 14.7)

- A break is a change in the probability distribution of the data
  - A change in the mean, variance, regression coefficients, etc
- It could be abrupt or gradual
- It often arises from changes in monetary policy
  - Central bank changes its inflation target or reaction function to macro shocks

## Problems caused by breaks

- They destroy external validity of time series models
- Thus, they are the most important cause of **forecast failure**
- They cause in-sample estimates of coefficients to be biased
- When coefficients have causal interpretation (see dynamic causal effects later), OLS estimates "average value" which does not correspond to the true causal effect at any period
- Hence, breaks can lead to incorrect interpretation of historical episodes and wrong policy prescriptions

## Testing for Breaks

- If the break date is known (meaning, if you have a specific hypothesis about when a change occurred), then we can test the null hypothesis of "No break" using a **standard F test**
- This F test is often called the Chow test (after econometrician Greg Chow who proposed it in 1960)
- To test for break at time period $\tau$ in a regression
  1. Create a dummy variable $D_t = 1$ if $t>\tau$, or 0 otherwise
  2. Estimate $Y_t = \beta_0 + \beta_1 X_t + \gamma_0 D_t + \gamma_1 D_t X_t + u_t$
  3. Test $H_0$: $\gamma_0 = \gamma_1 = 0$ using F test
- The dummy variable splits the sample in two subsamples
- In practice, the break point $\tau$ is rarely "known"

## Testing for a break at unknown break point

- When $\tau$ (the break date) is unknown, we can treat it as a parameter and estimate it alongside the regression coefficients
- That's the idea behind the Quandt Likelihood Ratio (QLR) test for a break at unknown break date
- The QLR test statistic is the largest Chow F statistic one can get across all candidate break dates
- Accordingly, the critical value is (much) larger than F
- For technical reasons you cannot consider break dates too close to the beginning and end of the sample

## The QLR test

Here are the details:

- Let $F(\tau)$ denote the Chow F-statistic for a break at date $\tau$
- Set $\tau_0$ = (integer part of $0.15T$) and $\tau_1$ = (int. part of $0.85T$)
- QLR stat = $\max[F(\tau_0), F(\tau_{0+1}), \ldots, F(\tau_1)]$
- Critical values for this were derived by Andrews (1990) (30 whole years after Quandt proposed the test!)
- They are given in SW Table 14.6
- For example: when you test stability of 5 coefficients:

| Test\Level | 10% | 5% | 1% |
|---|---|---|---|
| QLR | 3.26 | 3.66 | 4.53 |
| $F(5,\infty)$ | 1.85 | 2.21 | 3.02 |

## ***Example***: Is the Phillips curve relationship stable?

- Recall ADL(4,4) model in $\Delta \mathit{Inf}$ and *Unrate*
- Suppose we extend it to

$$\Delta \mathit{Inf}_t = \beta_{0,t} + \beta_1 \Delta \mathit{Inf}_{t-1} + \ldots + \beta_4 \Delta \mathit{Inf}_{t-4}$$
$$+\gamma_{1,t}\, \mathit{Unrate}_{t-1} + \ldots + \gamma_{4,t}\, \mathit{Unrate}_{t-r} + u_t$$

- Test $\mathbf{H_0}$: $\beta_{0,t}, \gamma_{1,t}, \ldots, \gamma_{4,t}$ are constant for all $t$ against

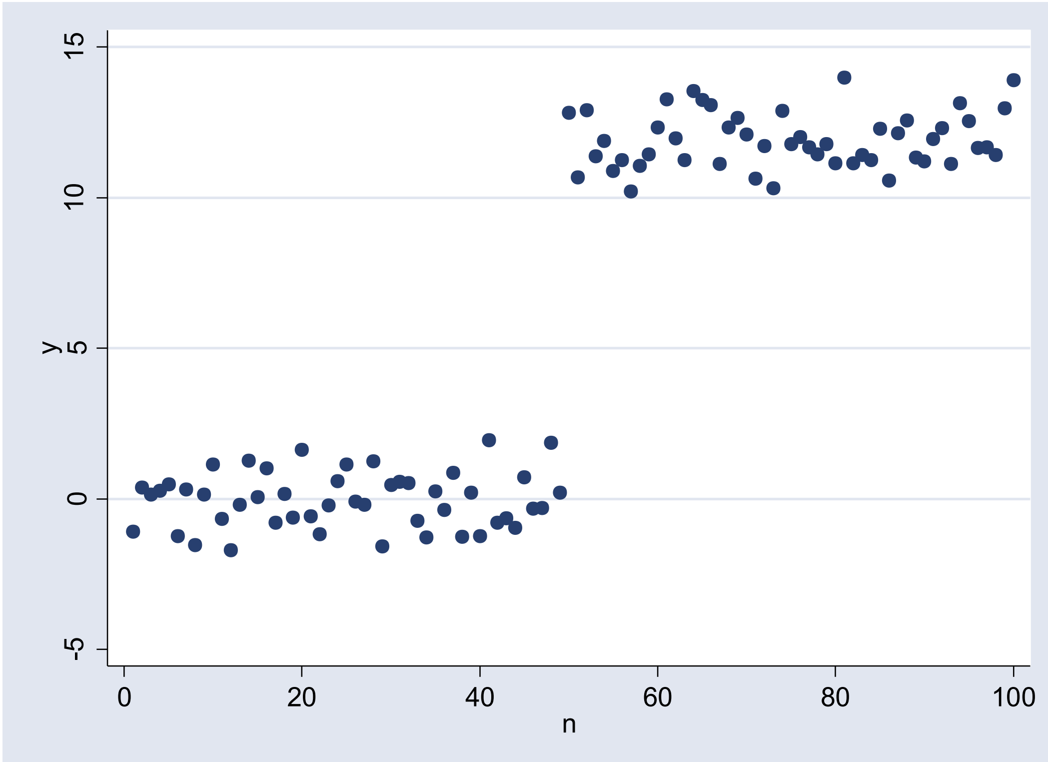
  $\mathbf{H_1}$: at least one of $\beta_{0,t}, \gamma_{1,t}, \ldots, \gamma_{4,t}$ changes at some $t = \tau$
- Number of restrictions $q = 5$
- QLR test = 5.16 > 4.53 (1% crit val of QLR test)
- Hence, reject $\mathbf{H_0}$ that Phillips curve has been stable
- Note: This does not tell us exactly how equation changes

- But if there is only one break, $\text{argmax}F(\tau)$ (the date at which largest Chow test occurs) is a good estimator for it

## Remedies for breaks

- Solution depends on the cause
- Extreme: if you know (or assume) the break dates, estimate separate (stable) models in different sub-periods
- Selective: if break is an artifact of outliers, remove those observations from the sample
- If break is the result of a regime shift (e.g., more active monetary policy after 1979, financial liberalization in the 1980s), you can model that shift using dummy variables
- Word of caution: It may be difficult to distinguish between multiple breaks and stochastic trends

- Consider the following time series




- This is stationary with a mean shift in middle of sample
- But if you ignore the shift and estimate AR(1) model, OLS estimate of AR(1) is 0.95 and DF test does not reject unit root hypothesis

# Cointegration
# (SW Section 16.4)

- Cointegration arises when two or more time series share a ***common stochastic trend***
- When this happens, regression analysis can reveal long-run relationships among time series
- But because the usual regression assumptions are violated, new methods are needed to analyse them
- Cointegration was discovered by econometricians Clive Granger and Robert Engle, who won the Nobel prize for it in 2003

# Definition of cointegration

- Suppose $Y_t$ and $X_t$ are integrated of order 1, denoted I(1)
- If for some coefficient $\theta$, $Y_t - \theta X_t$ is I(0), then $Y_t$ and $X_t$ are said to be **cointegrated**
- The coefficient $\theta$ is called the **cointegrating coefficient**
- If $Y_t$ and $X_t$ are cointegrated, then they have the same, or common, stochastic trend
- Computing $Y_t - \theta X_t$ eliminates this common stochastic trend
- This generalizes to multiple Xs: $Y_t - \theta_1 X_{1,t} - \ldots - \theta_k X_{k,t}$

# Economic interpretations of cointegration

- Cointegration can be thought of as a description of "long-run equilibrium"
  "[…] certain pairs of economic variables should not diverge from each other by too great an extent, at least in the long run. Thus, such variables may drift apart in the short run or according to seasonal factors, but if they continue to be too far apart in the long run, then economic forces, such as a market mechanism or government intervention, will begin to bring them together again." Granger (OBES, 1986)

- Cointegrating relation $Y_t - \theta X_t$ measures deviation ("error") from equilibrium
- A useful model of short run dynamics is the *Error Correction Model* (ECM)

$$\Delta Y_t = \beta_{10} + \beta_{11} \Delta Y_{t\text{-}1} + \ldots + \beta_{1p} \Delta Y_{t\text{-}p} + \gamma_{11} \Delta X_{t\text{-}1} + \ldots + \gamma_{1p} \Delta X_{t\text{-}p} + \alpha_1 (Y_{t-1} - \theta X_{t\text{-}1}) + u_t$$

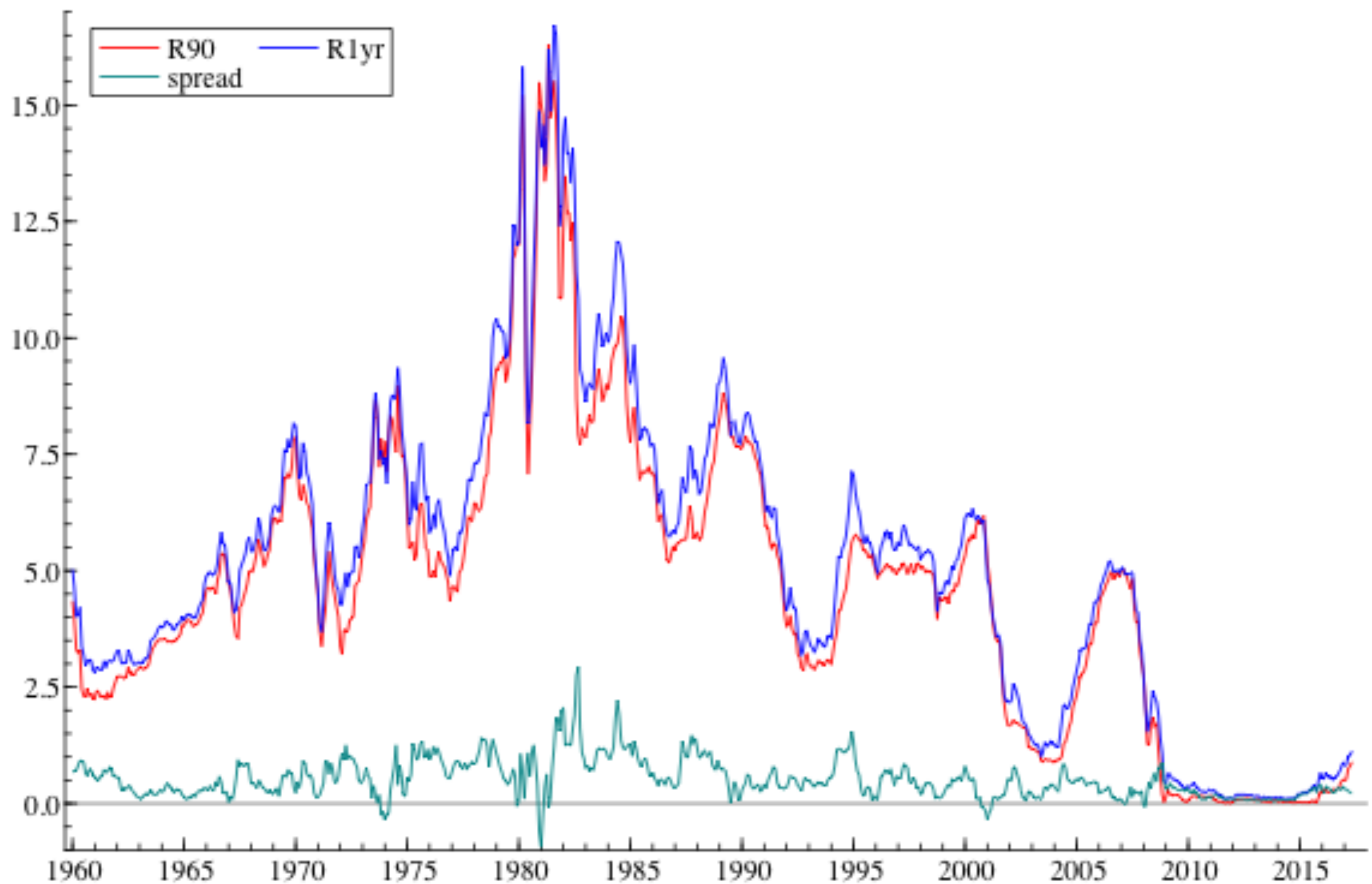
- This is ADL(p+1,p+1) in $Y_t$ and $X_t$
- We can write similar model for $X_t$ on $Y_t$
- According to this model, $Y_t$ and $X_t$ partially adjust to eliminate past errors
- ECM gives simple "causal" interpretation of movements in $Y_t$ and $X_t$

## *Example*: Expectations theory of the term structure

- Expectations theory of the term structure says that long rates (yields on long-maturity bonds) are the expected sum of short rates (yields on short-term treasury bills) over the same horizon
- So, if short rate has stochastic trend, then long rates must inherit the same trend (if the theory is right!)
- In other words, long and short rates must be cointegrated
- According to the theory, the cointegrating coefficient $\theta$ must be 1, i.e., the interest rate "spread" must be stationary

## *Example:* **Interest rates and spread -- OxMetrics**

- Load data InterestRates.xls and Plot R90, R1yr and spread = R1yr – R90



# *Example*: Interest rates and spread – OxMetrics ctd

- Perform ADF test using max 18 lags (1.5 years)
- Determine optimal lag by AIC and F

```
The sample is: 1962(1) - 2017(5) (684 observations and 3 variables)

R90: ADF tests (T=665, Constant; 5%=-2.87 1%=-3.44)
```

| D-lag | t-adf | beta Y_1 | sigma | t-DY_lag | t-prob | AIC | F-prob |
|---|---|---|---|---|---|---|---|
| 18 | -2.554 | 0.98763 | 0.3711 | 0.4061 | 0.6848 | -1.953 | |
| 17 | -2.526 | 0.98784 | 0.3709 | 0.8134 | 0.4163 | -1.956 | 0.6848 |
| 16 | -2.454 | 0.98826 | 0.3708 | 3.577 | 0.0004 | -1.958 | 0.6619 |
| 15 | -2.109 | 0.98986 | 0.3741 | -1.432 | 0.1526 | -1.941 | 0.0037 |
| 14 | -2.260 | 0.98918 | 0.3744 | 1.538 | 0.1246 | -1.941 | 0.0037 |
| 13 | -2.123 | 0.98987 | 0.3748 | 3.900 | 0.0001 | -1.940 | 0.0031 |
| 12 | -1.786 | 0.99142 | 0.3789 | -3.703 | 0.0002 | -1.920 | 0.0000 |
| 11 | -2.121 | 0.98975 | 0.3826 | 1.638 | 0.1018 | -1.902 | 0.0000 |
| 10 | -1.981 | 0.99045 | 0.3831 | -1.072 | 0.2841 | -1.901 | 0.0000 |
| 9 | -2.089 | 0.98997 | 0.3831 | 3.193 | 0.0015 | -1.902 | 0.0000 |
| 8 | -1.817 | 0.99125 | 0.3858 | 3.812 | 0.0002 | -1.890 | 0.0000 |
| 7 | -1.526 | 0.99260 | 0.3897 | 0.5100 | 0.6102 | -1.871 | 0.0000 |
| 6 | -1.494 | 0.99277 | 0.3895 | -6.483 | 0.0000 | -1.874 | 0.0000 |
| 5 | -2.018 | 0.98998 | 0.4015 | 1.428 | 0.1538 | -1.815 | 0.0000 |
| 4 | -1.903 | 0.99058 | 0.4018 | -0.2797 | 0.7798 | -1.815 | 0.0000 |
| 3 | -1.935 | 0.99046 | 0.4015 | 0.6304 | 0.5286 | -1.817 | 0.0000 |
| 2 | -1.889 | 0.99072 | 0.4013 | -5.521 | 0.0000 | -1.820 | 0.0000 |
| 1 | -2.405 | 0.98799 | 0.4102 | 9.328 | 0.0000 | -1.778 | 0.0000 |
| 0 | -1.619 | 0.99143 | 0.4360 | | | -1.657 | 0.0000 |

- Cannot reject $H_0$ that R90 has unit root at 5%

R1yr: ADF tests (T=665, Constant; 5%=-2.87 1%=-3.44)

| D-lag | t-adf | beta Y_1 | sigma | t-DY_lag | t-prob | AIC | F-prob |
|---|---|---|---|---|---|---|---|
| 18 | -2.196 | 0.99043 | 0.3589 | 0.8270 | 0.4085 | -2.020 | |
| 17 | -2.127 | 0.99078 | 0.3588 | 1.566 | 0.1178 | -2.022 | 0.4085 |
| 16 | -1.990 | 0.99140 | 0.3592 | 2.059 | 0.0399 | -2.021 | 0.2093 |
| 15 | -1.817 | 0.99215 | 0.3601 | -1.609 | 0.1080 | -2.018 | 0.0617 |
| 14 | -1.968 | 0.99152 | 0.3605 | 0.3604 | 0.7186 | -2.017 | 0.0417 |
| 13 | -1.945 | 0.99166 | 0.3603 | 2.507 | 0.0124 | -2.020 | 0.0734 |
| 12 | -1.737 | 0.99255 | 0.3617 | -3.319 | 0.0010 | -2.013 | 0.0122 |
| 11 | -2.034 | 0.99125 | 0.3645 | 3.922 | 0.0001 | -1.999 | 0.0003 |
| 10 | -1.707 | 0.99260 | 0.3685 | 0.5878 | 0.5569 | -1.979 | 0.0000 |
| 9 | -1.666 | 0.99280 | 0.3683 | -0.1512 | 0.8799 | -1.981 | 0.0000 |
| 8 | -1.685 | 0.99275 | 0.3680 | 3.797 | 0.0002 | -1.984 | 0.0000 |
| 7 | -1.406 | 0.99391 | 0.3718 | 0.5830 | 0.5601 | -1.966 | 0.0000 |
| 6 | -1.369 | 0.99408 | 0.3716 | -6.777 | 0.0000 | -1.968 | 0.0000 |
| 5 | -1.896 | 0.99156 | 0.3841 | 3.273 | 0.0011 | -1.903 | 0.0000 |
| 4 | -1.639 | 0.99267 | 0.3869 | -1.040 | 0.2989 | -1.890 | 0.0000 |
| 3 | -1.727 | 0.99230 | 0.3869 | 0.5450 | 0.5859 | -1.892 | 0.0000 |
| 2 | -1.690 | 0.99249 | 0.3867 | -6.585 | 0.0000 | -1.894 | 0.0000 |
| 1 | -2.269 | 0.98965 | 0.3989 | 10.17 | 0.0000 | -1.834 | 0.0000 |
| 0 | -1.456 | 0.99289 | 0.4286 | | | -1.692 | 0.0000 |

- Cannot reject $H_0$ that R1yr has unit root at 5%

spread: ADF tests (T=665, Constant; 5%=-2.87 1%=-3.44)

| D-lag | t-adf | beta Y_1 | sigma | t-DY_lag | t-prob | AIC | F-prob |
|---|---|---|---|---|---|---|---|
| 18 | -3.754** | 0.91438 | 0.1802 | -0.04717 | 0.9624 | -3.398 | |
| 17 | -3.807** | 0.91422 | 0.1800 | -1.635 | 0.1024 | -3.401 | 0.9624 |
| 16 | -4.117** | 0.90831 | 0.1803 | 0.3026 | 0.7623 | -3.400 | 0.2635 |
| 15 | -4.126** | 0.90940 | 0.1801 | -0.9105 | 0.3629 | -3.403 | 0.4300 |
| 14 | -4.345** | 0.90599 | 0.1801 | -0.08065 | 0.9357 | -3.405 | 0.4645 |
| 13 | -4.429** | 0.90569 | 0.1800 | 0.8495 | 0.3959 | -3.408 | 0.6086 |
| 12 | -4.349** | 0.90875 | 0.1799 | 0.2368 | 0.8129 | -3.410 | 0.6336 |
| 11 | -4.376** | 0.90960 | 0.1798 | 2.514 | 0.0122 | -3.412 | 0.7354 |
| 10 | -4.014** | 0.91777 | 0.1805 | -1.252 | 0.2108 | -3.406 | 0.2233 |
| 9 | -4.283** | 0.91347 | 0.1806 | -0.1327 | 0.8945 | -3.406 | 0.2023 |
| 8 | -4.372** | 0.91302 | 0.1805 | -0.2291 | 0.8188 | -3.409 | 0.2702 |
| 7 | -4.484** | 0.91222 | 0.1803 | -1.090 | 0.2761 | -3.412 | 0.3425 |
| 6 | -4.765** | 0.90830 | 0.1804 | -1.057 | 0.2910 | -3.414 | 0.3356 |
| 5 | -5.068** | 0.90433 | 0.1804 | -0.2209 | 0.8252 | -3.415 | 0.3338 |
| 4 | -5.221** | 0.90350 | 0.1803 | -2.750 | 0.0061 | -3.418 | 0.4027 |
| 3 | -5.974** | 0.89194 | 0.1811 | 2.506 | 0.0124 | -3.409 | 0.1055 |
| 2 | -5.548** | 0.90154 | 0.1819 | -4.722 | 0.0000 | -3.403 | 0.0293 |
| 1 | -6.906** | 0.87975 | 0.1848 | 4.381 | 0.0000 | -3.373 | 0.0000 |
| 0 | -5.986** | 0.89713 | 0.1873 | | | -3.347 | 0.0000 |

- Can reject $H_0$ that R1yr – R90 has unit root at 1%
- We can conclude that R1yr and R90 are **cointegrated**
- The cointegrating coefficient is 1

## Testing for cointegration when cointegrating coefficient is unknown

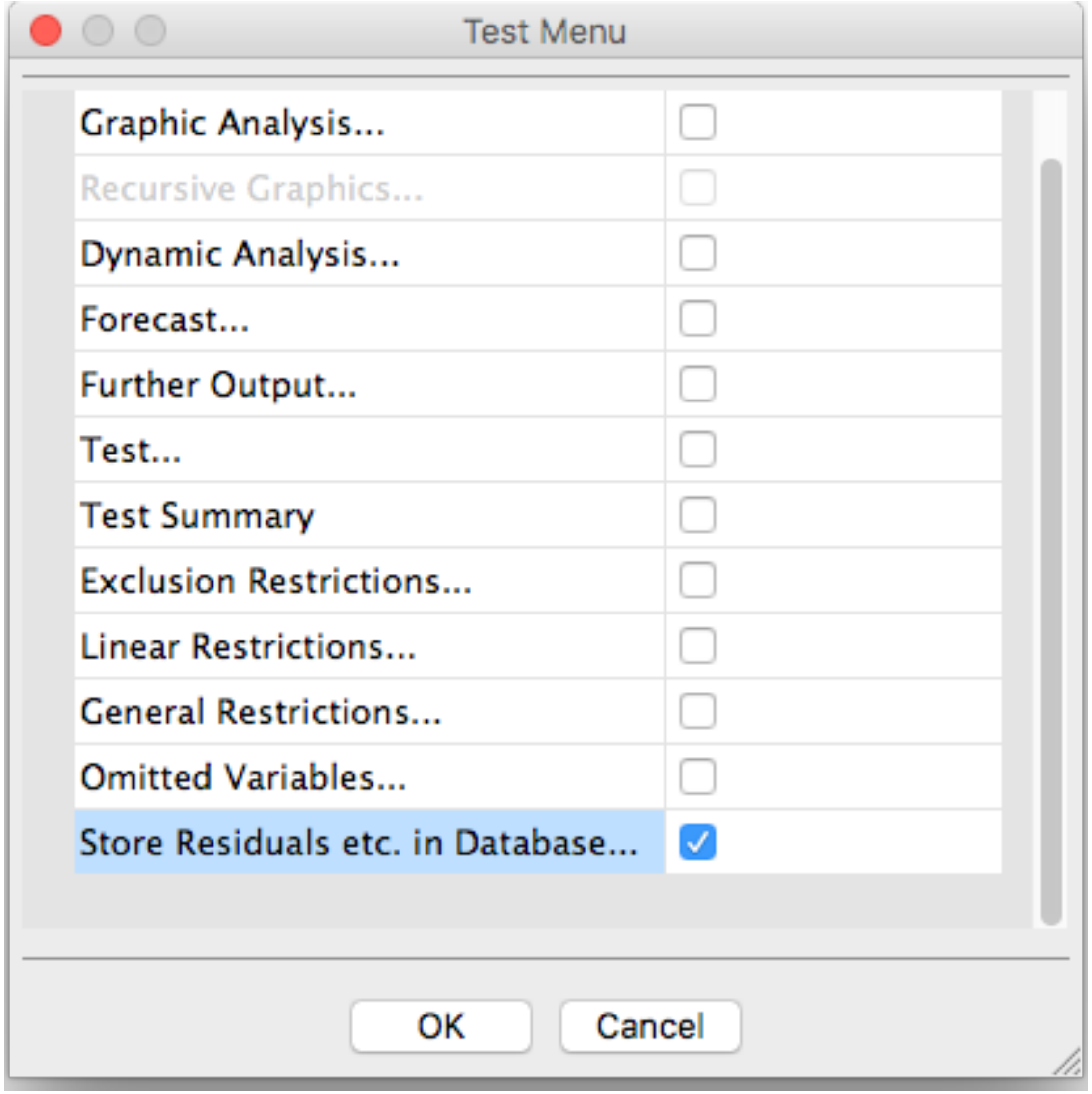
- In the previous example, economic theory suggested a value for the cointegrating coefficient: $\theta = 1$
  - Many other examples in economics, e.g., PPP, UIP
- When $\theta$ is unknown, we need to estimate it first
- It turns out OLS regression of Y on X yields a (super-) consistent estimator of $\theta$
- This motivates the so-called EG-ADF test for cointegration (due to Engle and Granger)
  1. Estimate static regression of $Y_t$ on $X_t$ and store residuals $\hat{u}_t$
  2. Perform ADF test on $\hat{u}_t$ using *different* critical values (to account for sampling uncertainty in estimating $\theta$)

## ***Example:*** Testing cointegration in interest rates

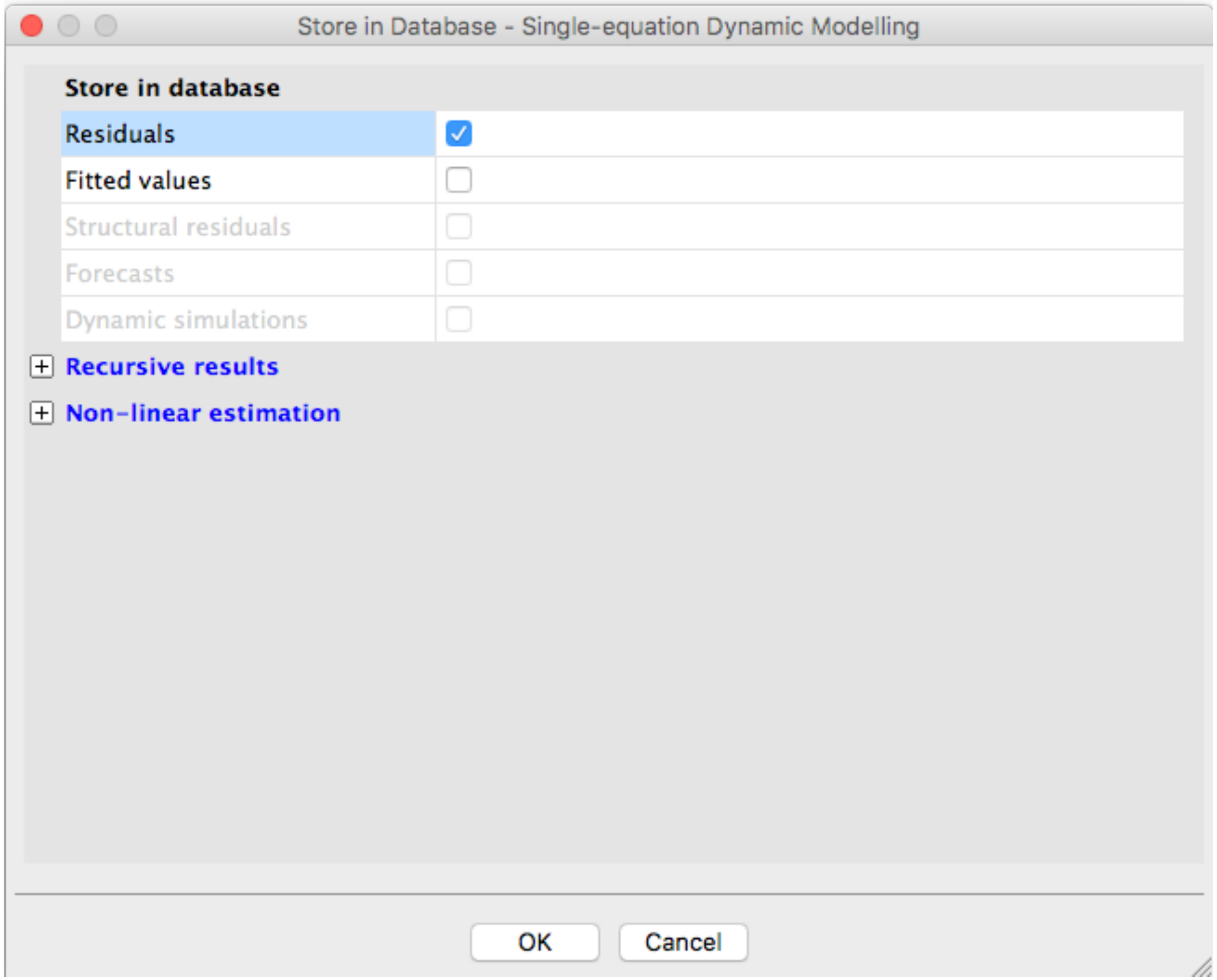
- Estimating static regression of R1yr on R90 yields

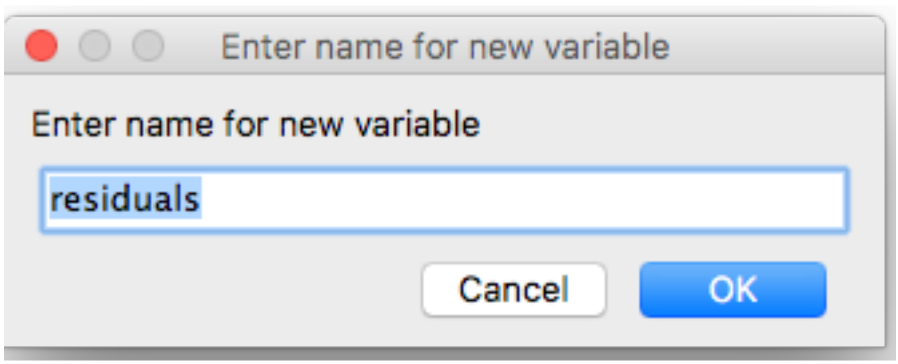
```
EQ( 1) Modelling R1yr by OLS
       The dataset is: /Users/sophocles/Dropbox/Teaching/Oxford/Depar
       The estimation sample is: 1960(1) - 2017(5)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
Constant             0.259903    0.02566     10.1  0.0000   0.1299
R90                   1.05561   0.004555     232.  0.0000   0.9874

sigma                0.378642  RSS               98.4947842
R^2                  0.987368  F(1,687) =   5.37e+04 [0.000]**
Adj.R^2               0.98735  log-likelihood      -307.514
no. of observations       689  no. of parameters          2
mean(R1yr)            5.17727  se(R1yr)             3.36651
```

- To store residuals, go to “Model\Test” and select “Store Residuals etc. in Database…”




Hit “OK”, select “Residuals”, hit “OK” and give the new series a name and hit “OK”


Store in Database - Single-equation Dynamic Modelling
Store in database
Residuals
Fitted values
Structural residuals
Forecasts
Dynamic simulations
Recursive results
Non-linear estimation
OK
Cancel


Enter name for new variable
Enter name for new variable
residuals
Cancel
OK

- Performing ADF test on the residuals yields

residuals: ADF tests (T=670, Constant; 5%=-2.87 1%=-3.44)

| D-lag | t-adf | beta Y_1 | sigma | t-DY_lag | t-prob | AIC | F-prob |
|---|---|---|---|---|---|---|---|
| 18 | -5.894** | 0.81555 | 0.1849 | 1.123 | 0.2617 | -3.346 | |
| 17 | -5.787** | 0.82319 | 0.1850 | -0.5332 | 0.5941 | -3.347 | 0.2617 |
| 16 | -6.071** | 0.81948 | 0.1849 | 1.246 | 0.2131 | -3.350 | 0.4619 |
| 15 | -5.940** | 0.82771 | 0.1849 | -0.2150 | 0.8298 | -3.350 | 0.3773 |
| 14 | -6.158** | 0.82628 | 0.1848 | 1.161 | 0.2459 | -3.353 | 0.5343 |
| 13 | -6.050** | 0.83365 | 0.1849 | 1.597 | 0.1107 | -3.354 | 0.4816 |
| 12 | -5.836** | 0.84318 | 0.1851 | 0.3571 | 0.7212 | -3.353 | 0.3186 |
| 11 | -5.906** | 0.84529 | 0.1850 | 2.642 | 0.0084 | -3.356 | 0.4127 |
| 10 | -5.456** | 0.85948 | 0.1858 | -0.7819 | 0.4346 | -3.349 | 0.0801 |
| 9 | -5.765** | 0.85511 | 0.1857 | 1.236 | 0.2167 | -3.351 | 0.1000 |
| 8 | -5.629** | 0.86167 | 0.1858 | 0.3555 | 0.7223 | -3.351 | 0.0938 |
| 7 | -5.687** | 0.86354 | 0.1857 | -0.6540 | 0.5134 | -3.354 | 0.1287 |
| 6 | -5.990** | 0.86001 | 0.1856 | -0.8545 | 0.3931 | -3.356 | 0.1580 |
| 5 | -6.381** | 0.85523 | 0.1856 | 0.08099 | 0.9355 | -3.358 | 0.1767 |
| 4 | -6.567** | 0.85569 | 0.1854 | -1.811 | 0.0706 | -3.361 | 0.2287 |
| 3 | -7.325** | 0.84484 | 0.1857 | 3.188 | 0.0015 | -3.359 | 0.1438 |
| 2 | -6.697** | 0.86170 | 0.1870 | -3.998 | 0.0001 | -3.347 | 0.0144 |
| 1 | -8.196** | 0.83681 | 0.1891 | 4.925 | 0.0000 | -3.326 | 0.0002 |
| 0 | -7.038** | 0.86238 | 0.1924 | | | -3.294 | 0.0000 |

- EG-ADF critical values depend on number of RHS regressors, see SW Table 16.2

- In this case, there is one regressor, so 1% critical value is -3.96 (smaller than the DF critical value -3.44)
- Since ADF test is -6.567 (AIC) or -7.325 (10% F) < -3.96 we reject H0 that R1yr and R90 are *not* cointegrated and we conclude that they *are* cointegrated

# Summary: Nonstationarity and cointegration

- Nonstationarity arises due to trends or breaks
- Trends can be deterministic or stochastic
- We can test for stochastic trends using the ADF test
- Order of integration of a time series is # of times it needs to be differenced to remove stochastic trends
- Regressions in trending variables can be *spurious*
- …unless the variables are *cointegrated*
- We test for cointegration using ADF or EG-ADF tests
- Breaks can destroy both internal and external validity of time series models
- We can test for breaks using the Chow or QLR tests