

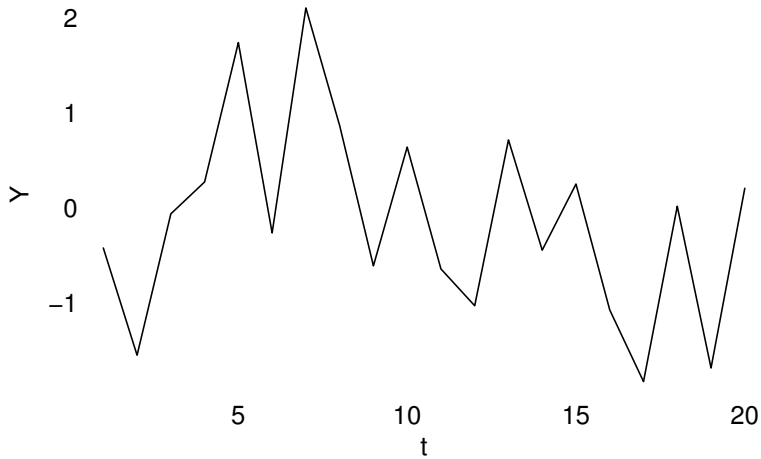
# Time-series data

School of Economics, University College Dublin

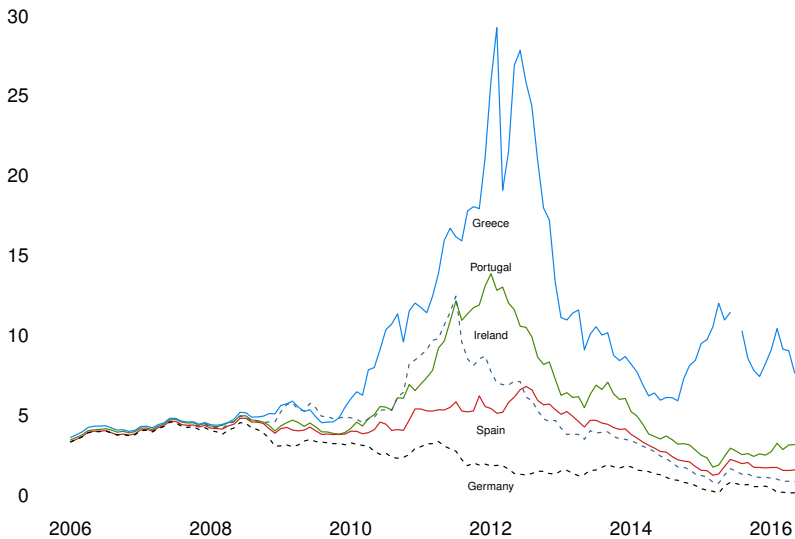
Spring 2018

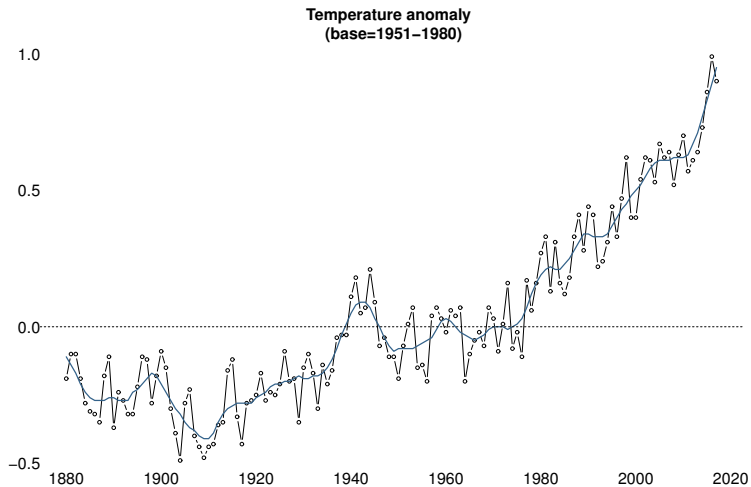
Things to do in empirical macroeconomics:

1. Data description: Describe and summarize macroeconomic data
2. Forecasting: Make macroeconomic forecasts
3. Structural inference: Quantify what we know and don't know about the true structure of the macro economy
4. Policy analysis: Advise on macroeconomic policy

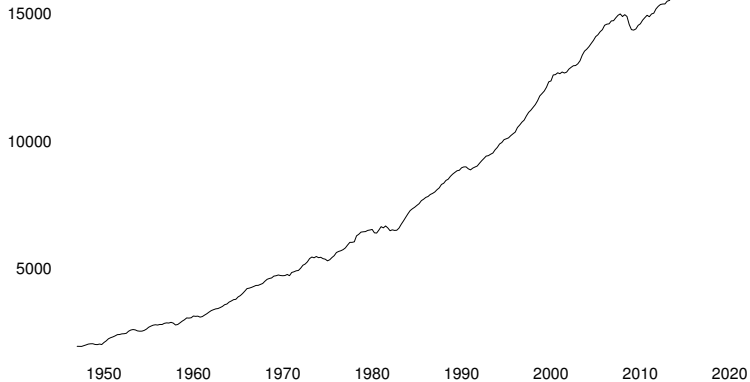


**Bond yields (percentage)**





**US real gdp  
(in billions)**



In general, time-series  $Y_t$  for  $t = 1, 2, \dots, T$  consists of

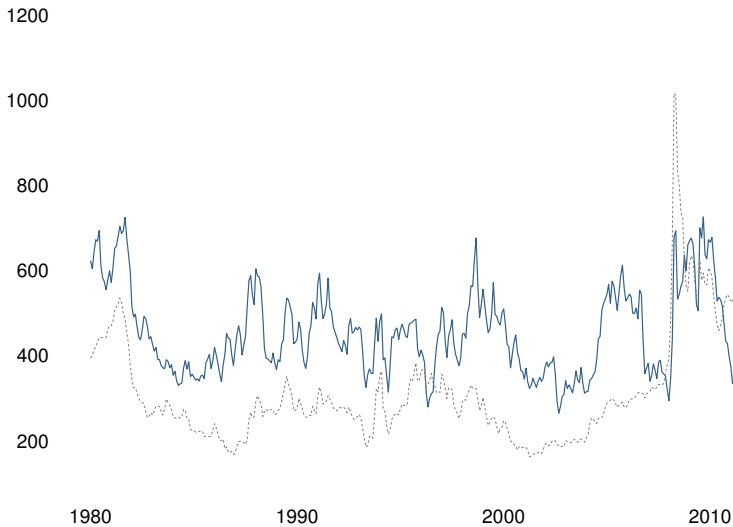
1. Trend component  $\tau_t$
2. Cyclical component  $c_t$
3. Error component  $\epsilon_t$

$$y_t = \tau_t + c_t + \epsilon_t \quad (1)$$

What we are often interested in are the short-term fluctuations such as business cycles which means that the data has to be broken down into

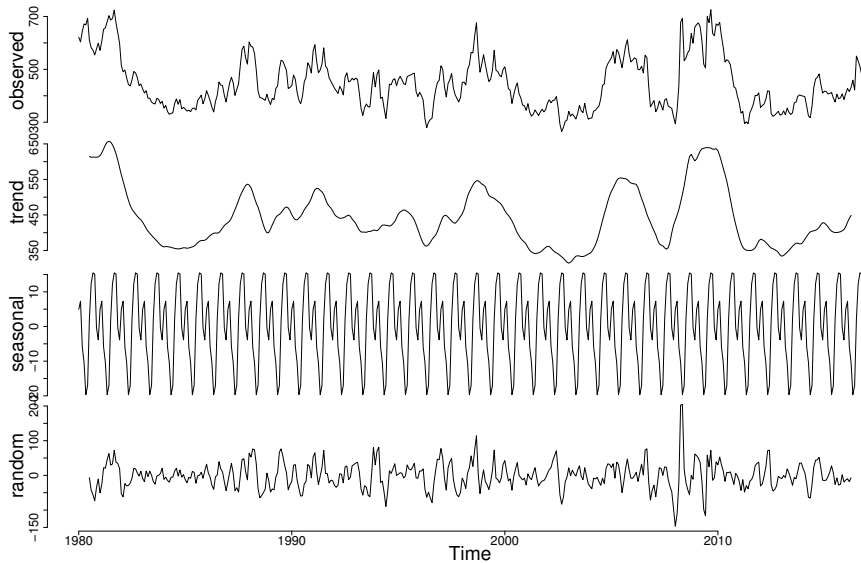
- ▶ A non-stationary long-run trend,
- ▶ A stationary cyclical component

# International rice prices (real)





# Decomposition of additive time series



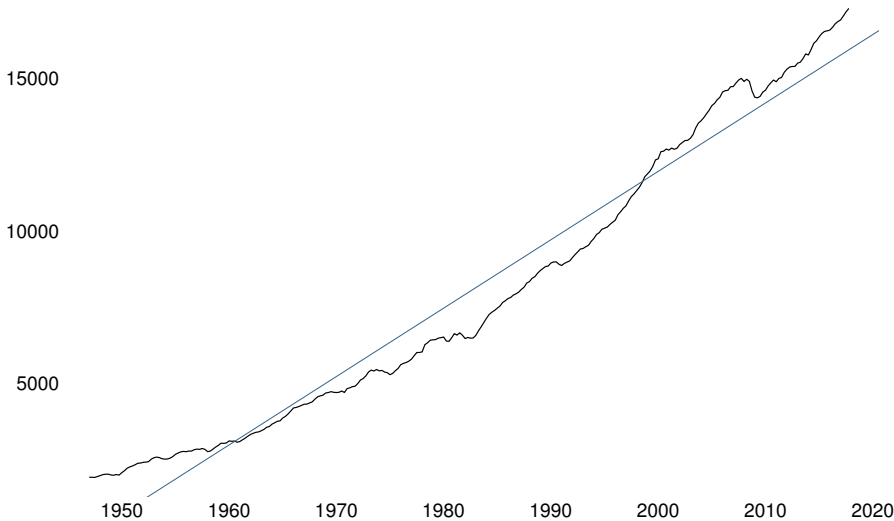
One of the simplest ways of detrending time-series data is using a log-linear model

$$\log Y_t = y_t = \alpha + gt + \epsilon_t \quad (2)$$

$\alpha + gt$  is the trend component

$\epsilon$  is the stationary cyclical component (with zero mean)

**US real gdp  
(in billions)**



# Linearly detrended

0.10  
0.05  
0.00  
-0.05  
-0.10  
-0.15

1950

1960

1970

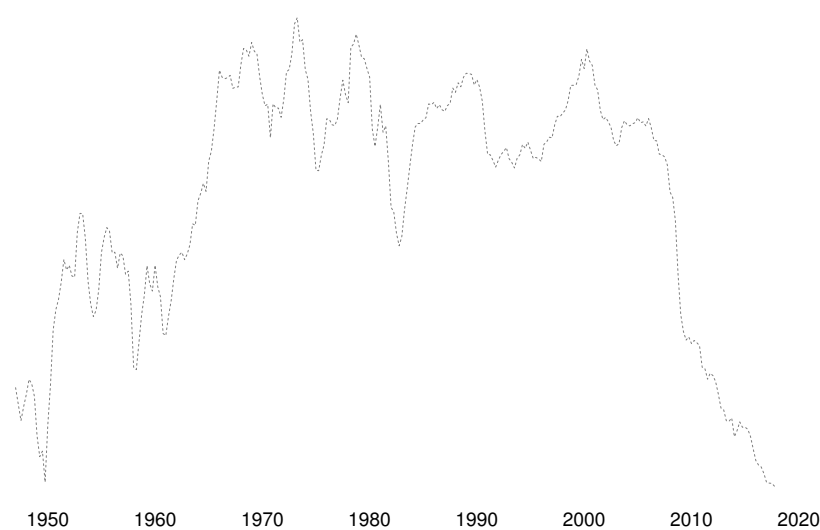
1980

1990

2000

2010

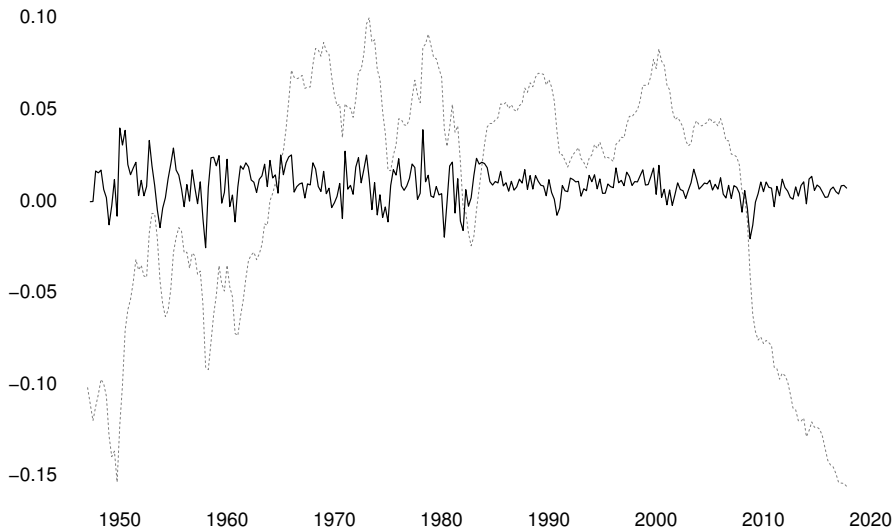
2020



Taking first-differences of log-transformed data will give something equivalent to the growth rate  $\Delta y_t$  which comprises of

1. Constant trend growth  $g$
2. Change in cyclical component  $\Delta \epsilon_t$

**Linearly detrended vs. First-differences**



For ease of notation let  $y_t$  be our time-series, can write growth as

$$\Delta y_t = \frac{y_t - y_{t-1}}{y_{t-1}} \quad (3)$$

Taking logs of growth rate gives

$$\Delta y_t = \log y_t - \log y_{t-1} \quad (4)$$

$$\begin{aligned}
 \log y_t - \log y_{t-1} &= \log \left( \frac{y_t}{y_{t-1}} \right) \\
 &\approx \frac{y_t}{y_{t-1}} - 1 \\
 &= \frac{y_t}{y_{t-1}} - \frac{y_{t-1}}{y_{t-1}} \\
 &= \frac{y_t - y_{t-1}}{y_{t-1}}
 \end{aligned}
 \tag{5}$$

**NB-** If  $y$  is close to 1, this means that  $\log y$  will be close to  $y - 1$ , and  $\frac{y_t}{y_{t-1}}$  is likely close to 1



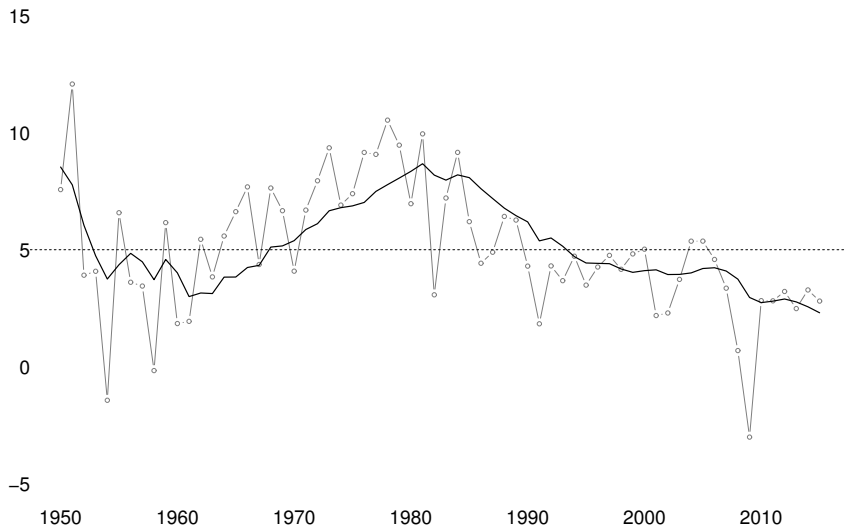
Although straightforward, there is a caveat using a straight line to detrend the data. Suppose that the correct model is

$$y_t = g + y_{t-1} + \epsilon_t \quad (6)$$

Recall that growth has a constant component  $g$  and random component  $\epsilon_t$ : two important implications

1. Cycles are the accumulation of all the random shock over time that affected  $\Delta y_t$
2. Expected growth rate will be  $g$ ; irrespective of what happened in the past

**NB**-Another issue is medium-run changes in  $g$



Since  $\Delta y_t$  is stationary, taking first-differences will remove non-stationary stochastic trend component ( $y_{t-1}$ ) in the data

$$\Delta y_t = y_t - y_{t-1} = g + \epsilon_t \quad (7)$$

Therefore, fitting a log-linear line to the data, there might appear to be a mean-reverting cyclical component, which is not there.

An example: data is randomly generate for following model

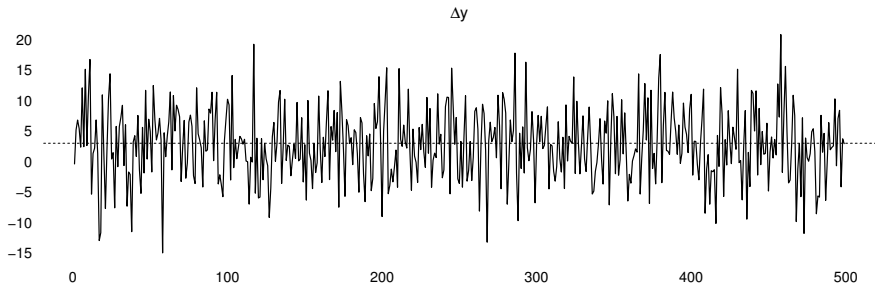
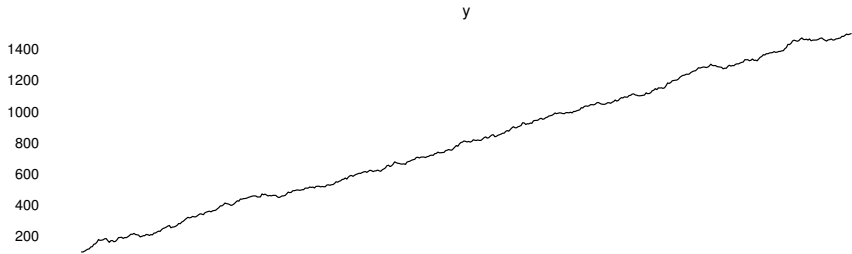
$$y_t = g + y_{t-1} + \epsilon_t \quad (8)$$

$g=3$ ; starting value for  $y$  is 100

$$\epsilon \sim N(0, 6)$$

Take first-differences

$$\Delta y = y_t - y_{t-1} \quad (9)$$



Incorrect usage of time-series dynamics can lead to errors in identification. An interesting example comes from a paper on the link between precipitation, economic growth, and conflict by Miguel et al. (2004) who use the following IV-2SLS model

$$\Delta Y_{it} = a_{1i} + b_1 X'_{it} + c_{1,0} \Delta R_{it} + c_{1,1} \Delta R_{it-1} + d_{1i} y_t + \epsilon_{1it} \quad (10)$$

$$C_{it} = \alpha_{2i} + \beta_2 X'_{it} + \gamma_{2,0} \Delta \hat{Y}_{it} + \gamma_{2,1} \Delta \hat{Y}_{it-1} + \delta_{2,1} + \epsilon_{2it} \quad (11)$$

One issue here is that precipitation is mean-reverting.

Ciccone (2011) argues the following; consider the following model to link conflict and precipitation

$$P(\text{conflict}_t) = a\Delta R_t + b\Delta R_{t-1} \quad (12)$$

Can write this in levels as

$$P(\text{conflict}_t) = \alpha_0 \log R_t + \alpha_1 \log R_{t-1} + \alpha_2 \log R_{t-2} \quad (13)$$

Given that

$$\Delta R_t = \log R_t - \log R_{t-1} \quad (14)$$

The estimates in Eq.12 are given by

$$a = \frac{2\alpha_0 - (\alpha_1 + \alpha_2)}{3}; b = \frac{(\alpha_0 + \alpha_1) - 2\alpha_2}{3} \quad (15)$$

i.e. parameters in eq.12 are a mixture of those in eq.13.



A commonly used method to detrend data is the Hodrick-Prescott (HP) filter; taking  $y_t$  the filter minimises

$$\sum_{t=1}^N [(y_t - y_t^*)^2 + \lambda(\Delta y_t^* - \Delta y_{t-1}^*)] \quad (16)$$

Here  $y^*$  is the time-varying trend;  $\lambda$  is a penalty parameter

- For quarterly data  $\lambda = 1,600$

The HP filter does two things:

1. Minimise the sum of squared deviations between output and its trend

$$(Y_t - Y_t^*)^2 \quad (17)$$

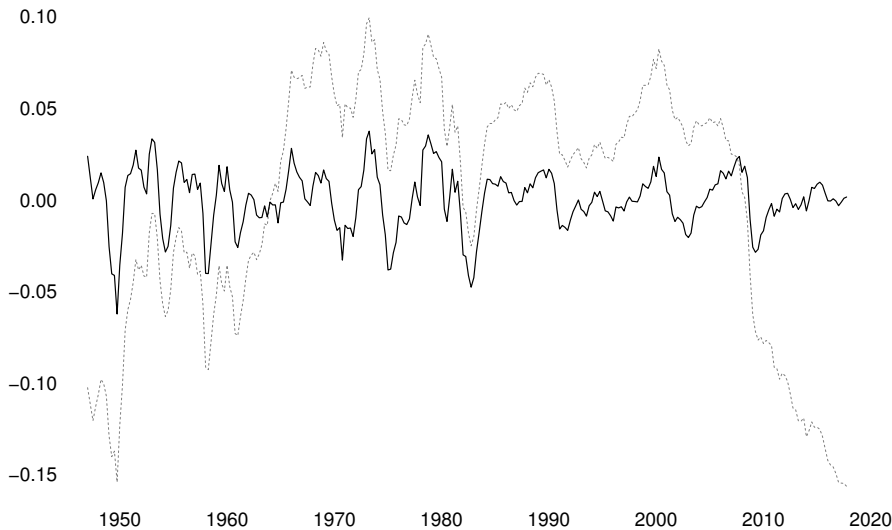
2. Minimising the change in the trend growth rate

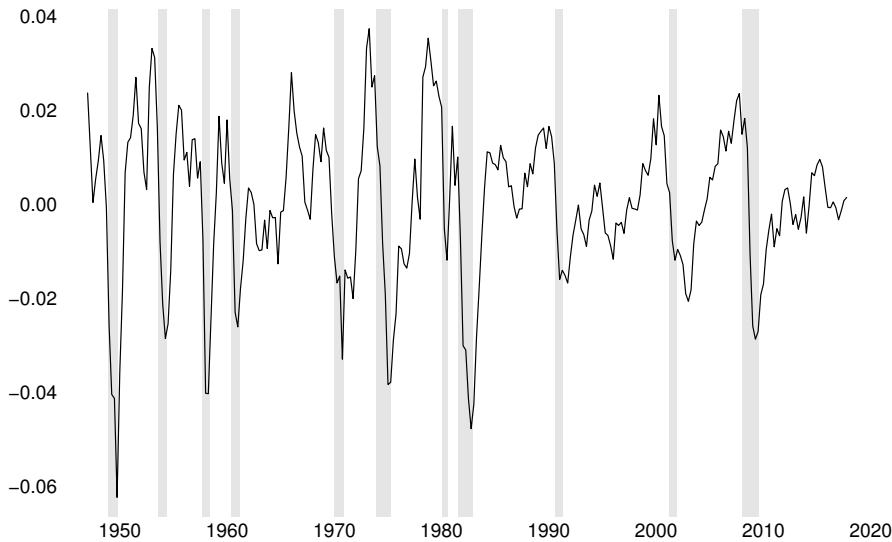
$$\lambda(\Delta Y_t^* - \Delta Y_{t-1}^*) \quad (18)$$

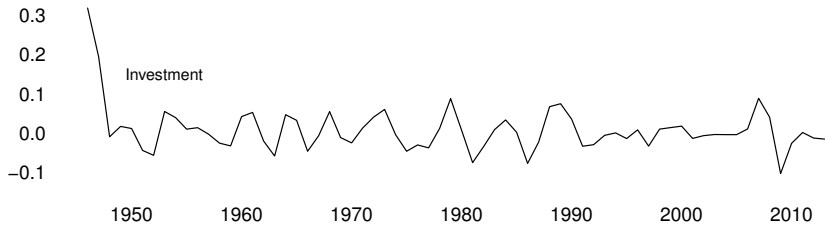
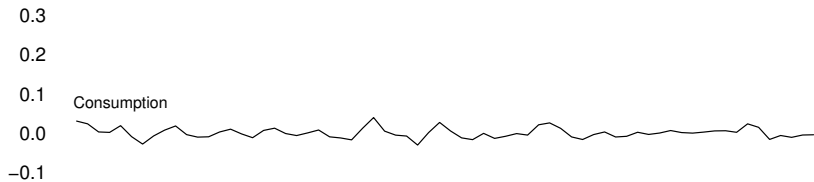
Important is that the HP-filter let's the growth rate vary over time.

- Larger  $\lambda$  means smoother trend

**Linearly detrended vs. Hodrick–Prescott filter**

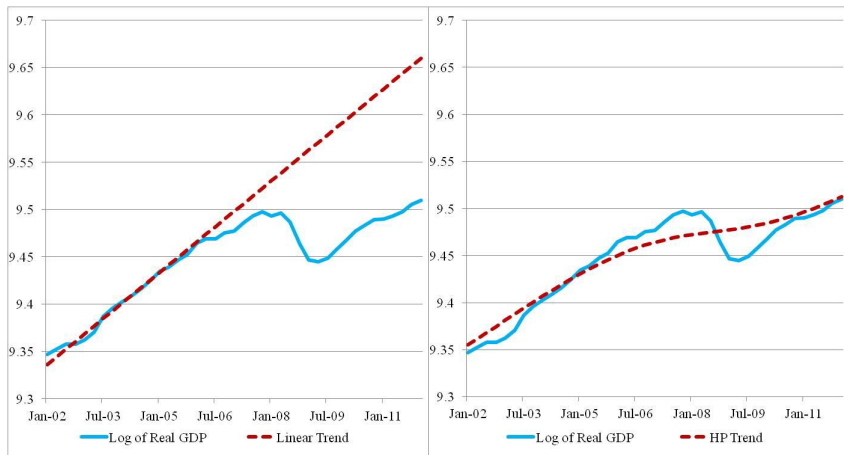






Can use the HP-filter to look at cycles in different components of GDP (this case UK): shows that consumption tends to be more stable

1. The consumption of specifically non-durable goods such as food is stable over time
2. People tend to spend a constant amount of money regardless of temporary shocks (permanent income hypothesis)



Source: James Bullard

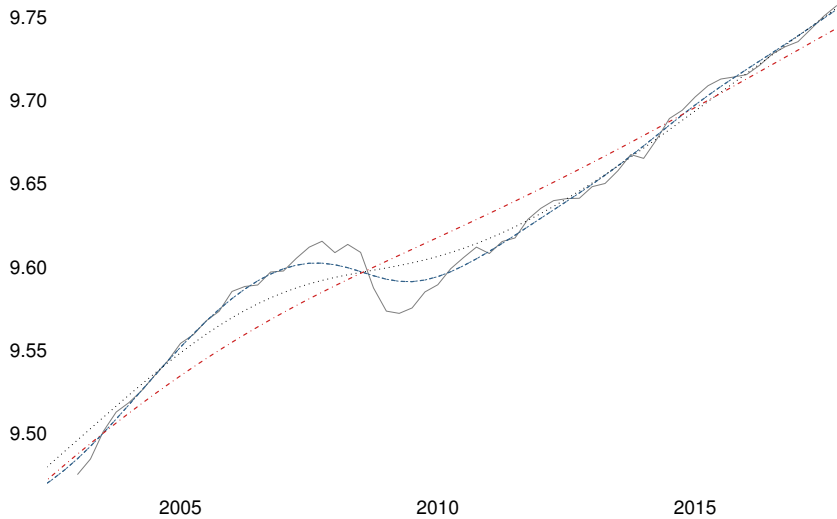
Statistical technique is only useful if assumptions reflect economic reality. But HP-filter assumes that deviations from trend are short term and correct themselves quickly

- ▶ Prolonged periods below potential GDP not possible

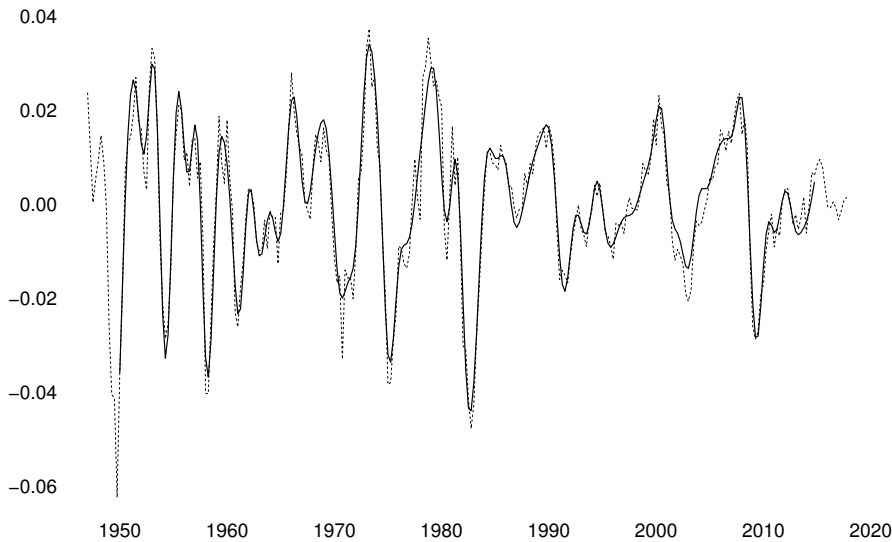
There is an endpoint problem in the filter

- ▶ Smoothed series close to observed data at beginning and end of period

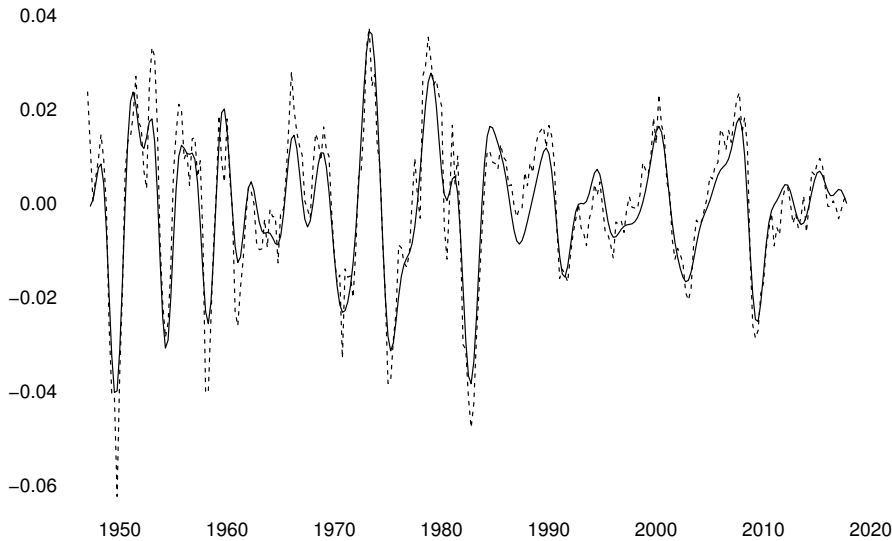




HP filter vs, Baxter-King filter



**HP filter vs. Wavelet filter**



The cyclical components in time-series data can be autocorrelated and exhibit random-looking fluctuations. Simple way to capture these dynamics is through an Autoregression (AR) model, e.g. AR(1)

$$y_t = \rho y_{t-1} + \epsilon_t \quad (19)$$

$\rho$  is propagation mechanism

- Determines at which speed a shock in  $\epsilon$  fade away

Time path of  $y$  after a shock in  $\epsilon$  is the Impulse Response Function (IRF), i.e.  $t$  follows

$$\epsilon_t + 1, \epsilon_{t+1}, \epsilon_{t+2}, \dots \quad (20)$$

Instead of

$$\epsilon_t, \epsilon_{t+1}, \epsilon_{t+2}, \dots \quad (21)$$

This means there is an incremental effect in all future periods of a unit shock today.

Imagine AR(1) series starting at 0 with shock  $\epsilon_t = 1$  followed by 0 shocks for  $t + 1$ . We get that for period  $t$

$$y_t = 1 \tag{22}$$

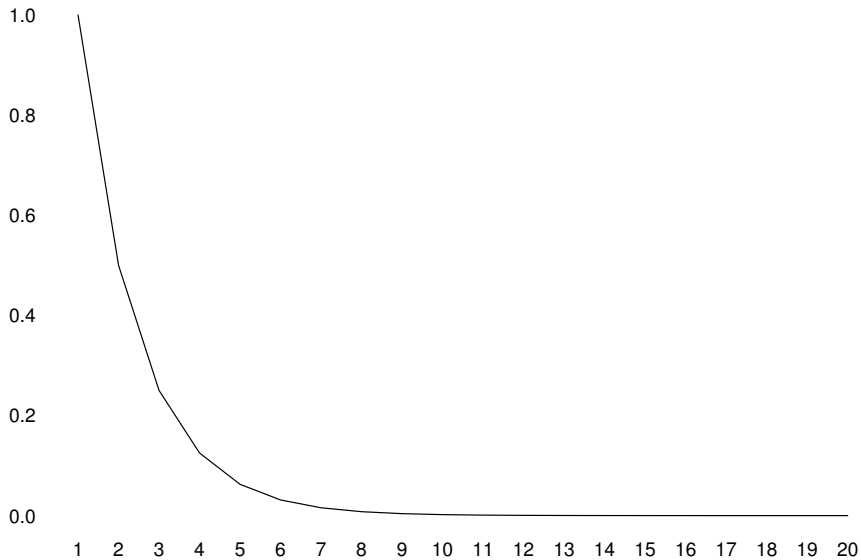
For period  $t + 1$

$$y_{t+1} = \rho \tag{23}$$

For period  $t + n$

$$y_{t+n} = \rho^n \tag{24}$$

**IRF AR(1)**



Let's consider volatility which is determined by

1. Size of shock in  $\epsilon$
2. Strength of propagation mechanism  $\rho$

So for  $y_t$  we get

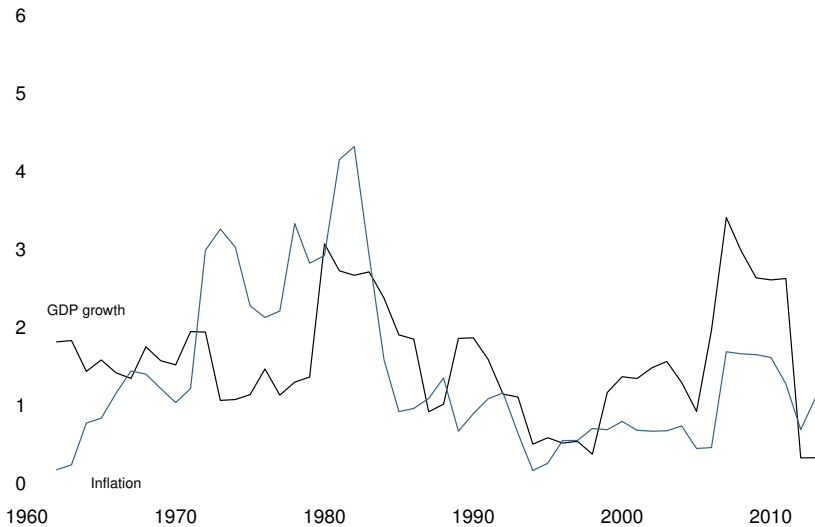
$$\begin{aligned}\sigma_y^2 &= \rho^2 \sigma_y^2 + \sigma_\epsilon^2 \\ &= \frac{\sigma_\epsilon^2}{1 - \rho^2}\end{aligned}\tag{25}$$

$\sigma_\epsilon^2$  is the variance of  $\epsilon_t$

The long run variance of  $y_t$  is the same as the long-run variance of  $y_{t-1}$



## Rolling 5-year average



**the Great Moderation:** Since mid-1980s output and inflation have become less volatile, which could be due to

1. Smaller shocks (lower values for  $\epsilon_t$ )
  - ▶ Less random policy shocks
  - ▶ Smaller shocks from goods and/or financial markets
  - ▶ Smaller supply shocks
2. Weaker propagation mechanisms (lower values for  $\rho$ )
  - ▶ Policy became more stabilizing
  - ▶ More stable economy
  - ▶ Stabilisation of economy due to financial modernisation

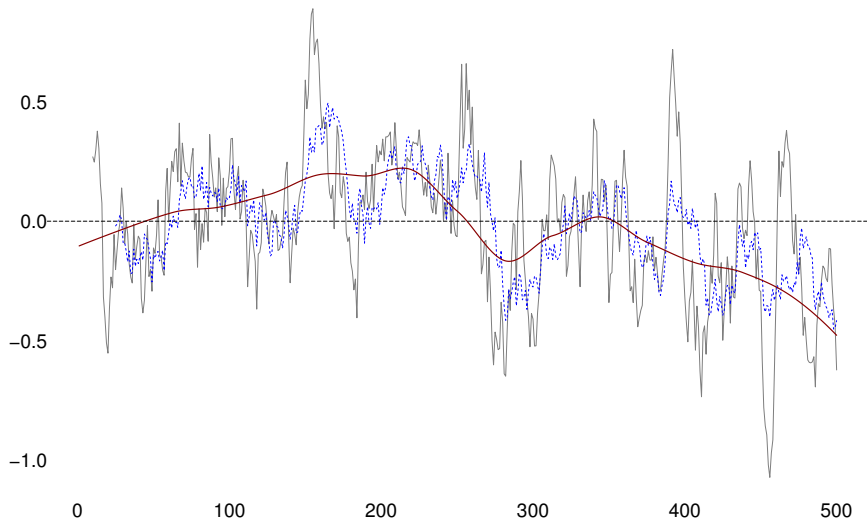
**Slutsky effect:** when data are uncorrelated smoothing can introduce appearance of irregular oscillations (Kelly & O'Grada, 2014). There are two different definitions, the formal one

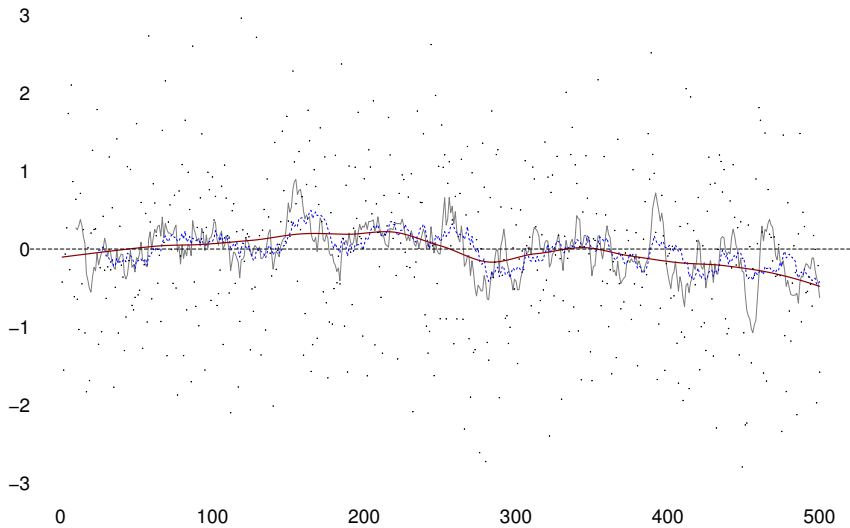
$$f(\omega) = \frac{1}{m^2} \frac{1 - \cos m\omega}{1 - \cos \omega} \quad (26)$$

This is the transfer function for a moving average of  $m$  periods.

And there is the colloquial one

*Applying a moving average to a white noise series will generate the appearance of irregular oscillations, as the filter is distorted by runs of high or low observations.*





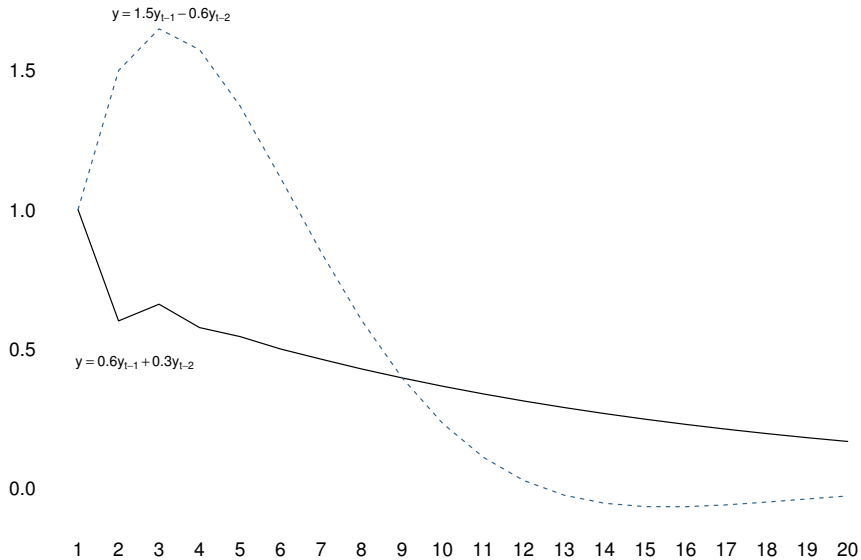
Often macroeconomic dynamics extend beyond AR(1) process; consider AR(2) model

$$y_t = \alpha + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \epsilon_t \quad (27)$$

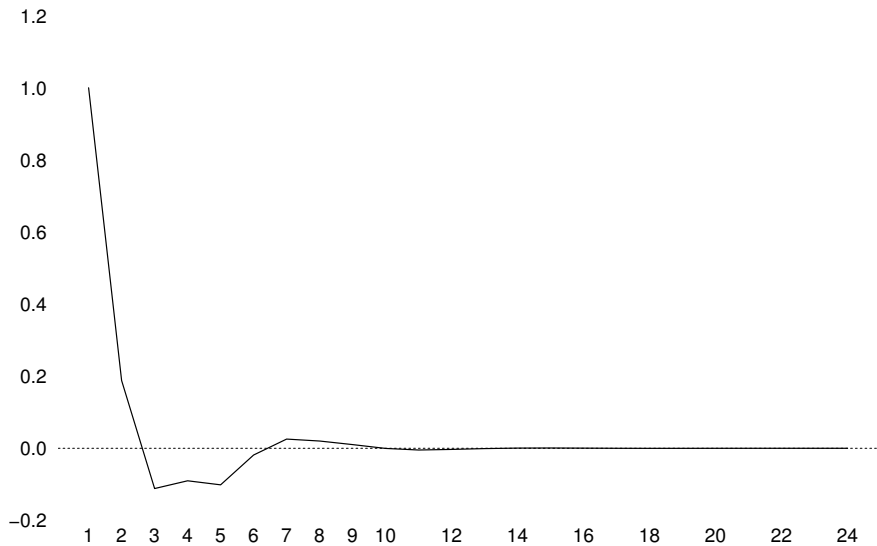
IRF can take on various forms based on  $\rho_1, \rho_2$  values

- ▶ More complex responses can be generate by AR(p) model
- ▶ Dynamic properties of model depend on number of lags  $p$  included in model

# IRF AR(2)



**AR(4) model for rice**





Concerning model notation; we can use the lag operator  $L$

- The operator moves the time-series back in time

So instead of

$$y_{t-1} \quad (28)$$

We can write

$$Ly_t \quad (29)$$

Similarly

$$L^2 y_t = y_{t-2} \quad (30)$$

Can use the lag operator in model specification when the model includes a number of lags. Can write model

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t \quad (31)$$

as

$$y_t = A(L)y_t + \epsilon_t, \quad A(L) = a_1 L + a_2 L^2 \quad (32)$$

Alternatively you can write it as

$$B(L)y_t = \epsilon_t, \quad B(L) = 1 - a_1 L + a_2 L^2 \quad (33)$$

AR models are useful in understanding the dynamics of individual variables

- ▶ But they ignore the relationships between variables

Vector Autoregressions (VAR) model the dynamics between  $n$  different variables, allowing each variable to depend on the lagged values of all the variables

- ▶ Can examine the impulse response of  $n$  variables to all  $n$  shocks

Simplest VAR model has two variables and one lag

$$y_{1t} = a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + \epsilon_{1t} \quad (34)$$

$$y_{2t} = a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + \epsilon_{2t} \quad (35)$$

Lot of talk about shocks: some sources include

1. Policy changes, which are not captured by the systematic component of the VAR equation
2. Changes in preferences, such as work vs. leisure or spending vs. saving
3. Technology shocks, which are random changes in the productivity of firms
4. Shocks to various frictions, like changes in the efficiency with which markets (labour, goods, financial) work

Time-series perspective is central to economic fluctuations in macroeconomics

- ▶ Cycles are determined by various random shocks, propagated throughout the economy over time

VARs are commonly used for modeling macroeconomic dynamics and the effect of shocks

- ▶ Can help explain *how* things work, but not why things work the way they do

Dynamics Stochastic General Equilibrium model build upon VAR framework, but dynamics are derived from economic theory

- ▶ In this framework agents are i) rational and ii) optimising