Modeling Unemployment Rates

Martin Burger
R PROGRAMMING TUTOR

www.r-tutorials.com

Trending Data

Understanding the concept of trending time series

- A clear pattern that most models can catch

Theoretical background

Data: Labor force participation rate of Spain

Holt exponential smoothing method

- Trend damping feature

ARIMA model

Comparison plot of all models with ‘autolayer’

Working with Trending Time Series

Working with Trending Time Series

Handling a time series with a trend

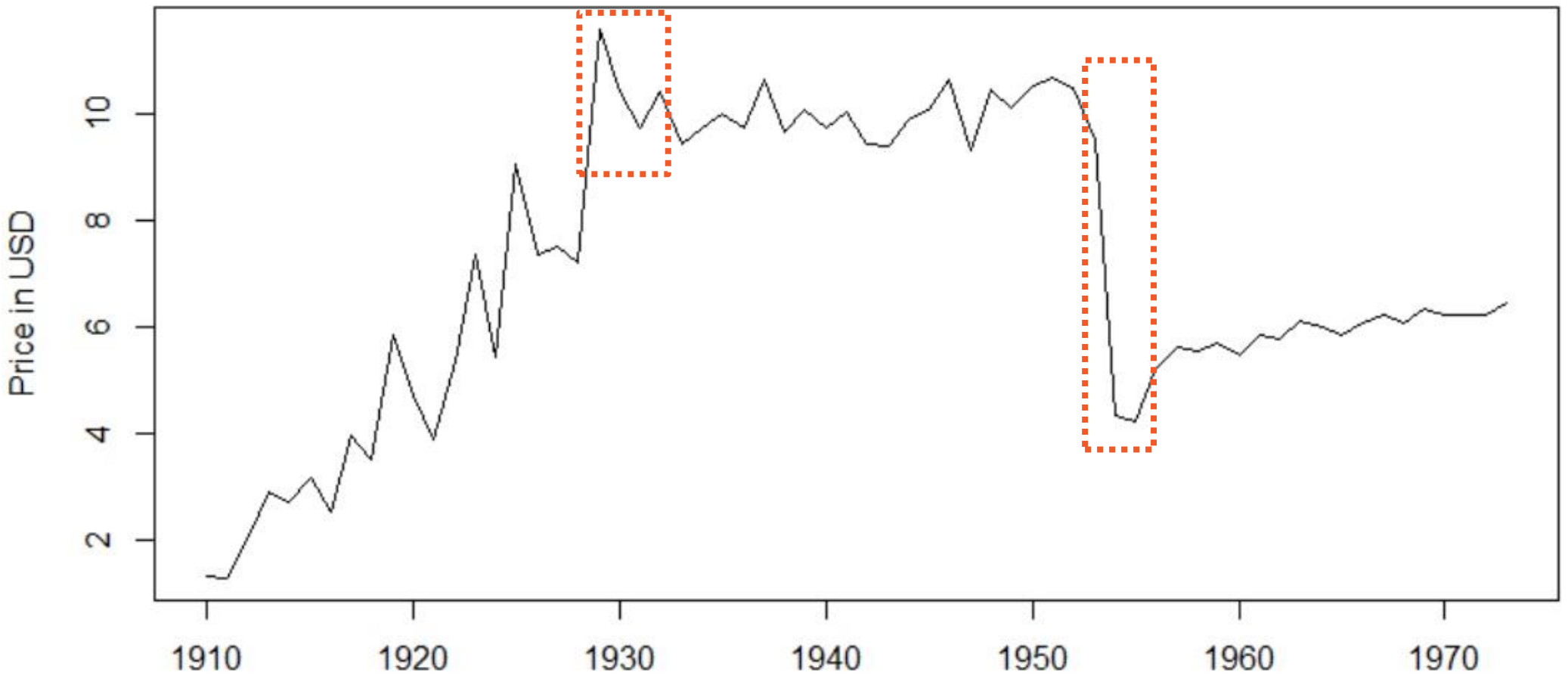
Available methods

Things to watch out for

Trend

A long-term pattern that gives a clear direction to a time series.

Trend with Changepoints



Changepoints in a Trend

Trends always come to an end

- Random point
- Inherent point

Inherent changepoints are predictable

Trend dampening likely occurs before an inherent changepoint

- Damping parameter

A trend is a clear pattern that makes a model more accurate or even enables it

Modeling Time Series with a Trend

ARIMA model

Exponential smoothing models

Labor Force Participation Rate of Spain

Measuring Employment of a Country

Unemployment Rate

The number of unemployed people as a percentage of the labor force

Common metric used in media and politics

Prone to manipulations

Who can contribute to the unemployment rate?

Labor Force Participation Rate

Labor force divided by the total of working-age population

Working age: 25-54

Less prone to manipulations

Labor force participation rate of Spain

Source: gapminder.org

Yearly data: 1980-2007

Data: lfpr_spain.csv

```
spain = ts(spain$x, start = 1980)
```

Converting the Data into a Time Series

Updating the existing object ‘spain’

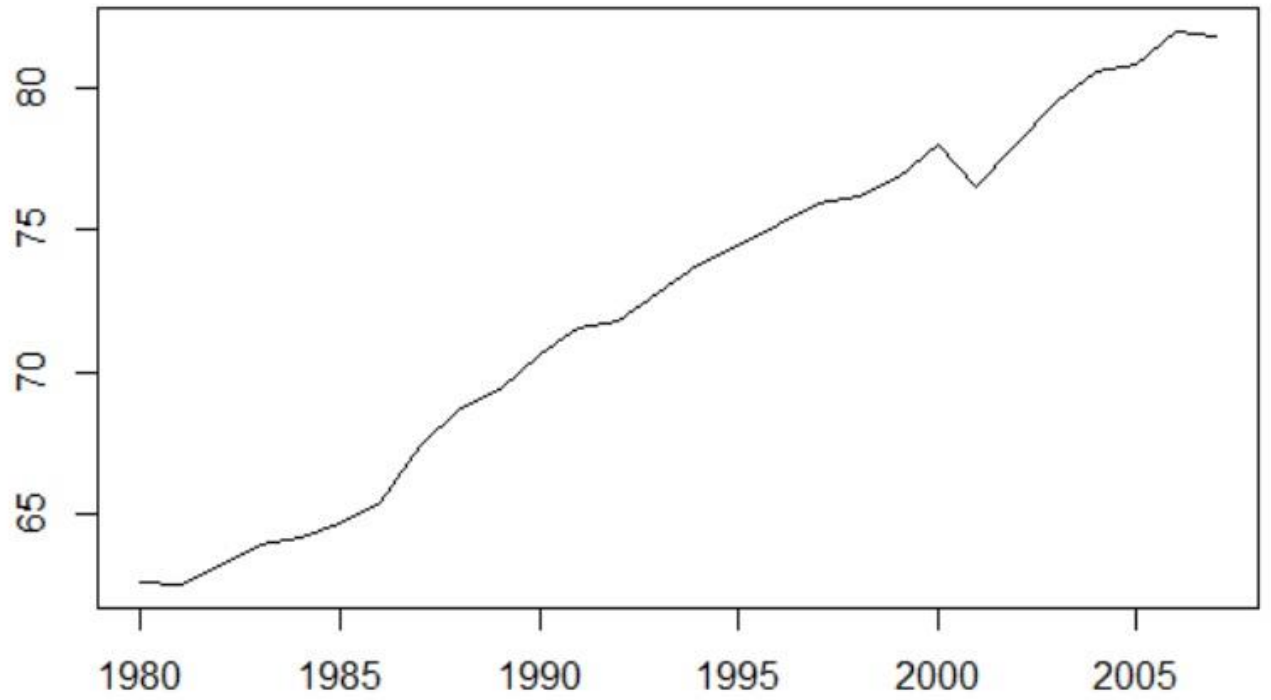
- New time series object takes only data from column ‘x’
- Timestamp starts in 1980
- No frequency argument is required

Trending data

ARIMA model

Exponential smoothing model

The rate cannot go higher than 100%



Exponential Smoothing for Trending Data

Exponential Smoothing in Library Forecast

`ses()`

Simple exponential smoothing

`holt()`

Linear trend model

`hw()`

Holt-Winters exponential smoothing

`ets()`

Automated exponential smoothing

Forecast Equation with Holt's Linear Trend Method

$$y_{t+h} = l_t + hb_t$$

Estimated value: y_{t+h}

Level (constant): l_t

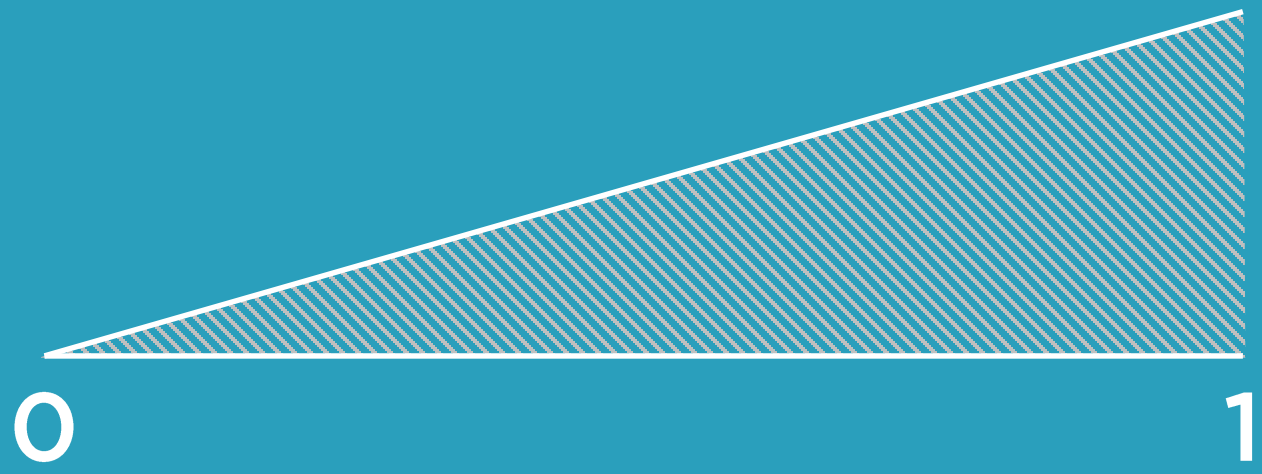
Trend value: b_t

Number of forecasting steps: h

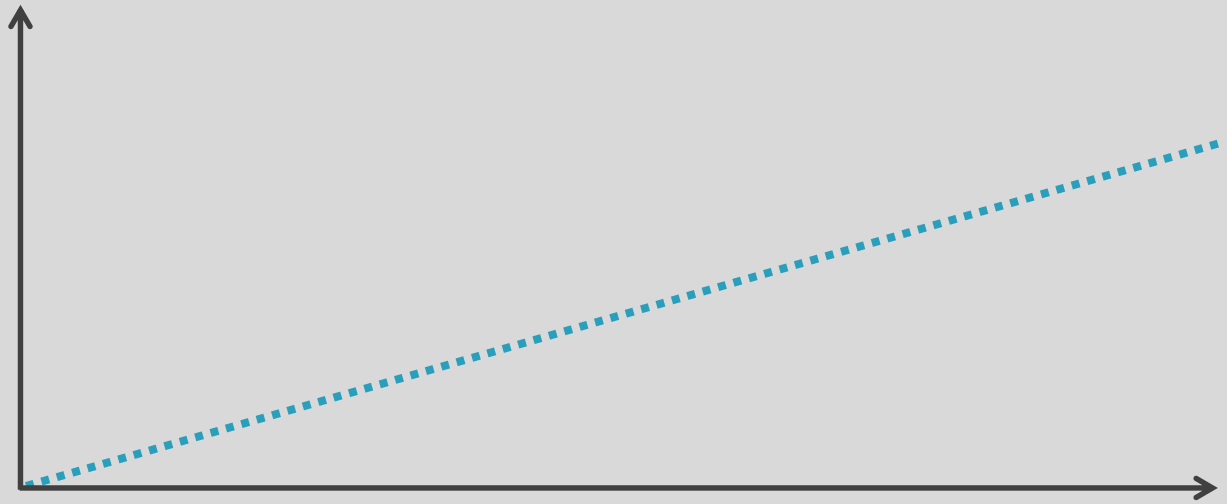
Reactiveness is adjusted by two smoothing parameters α and β

- Smoothing the level with α
- Smoothing the trend with β

Reactiveness of the Model

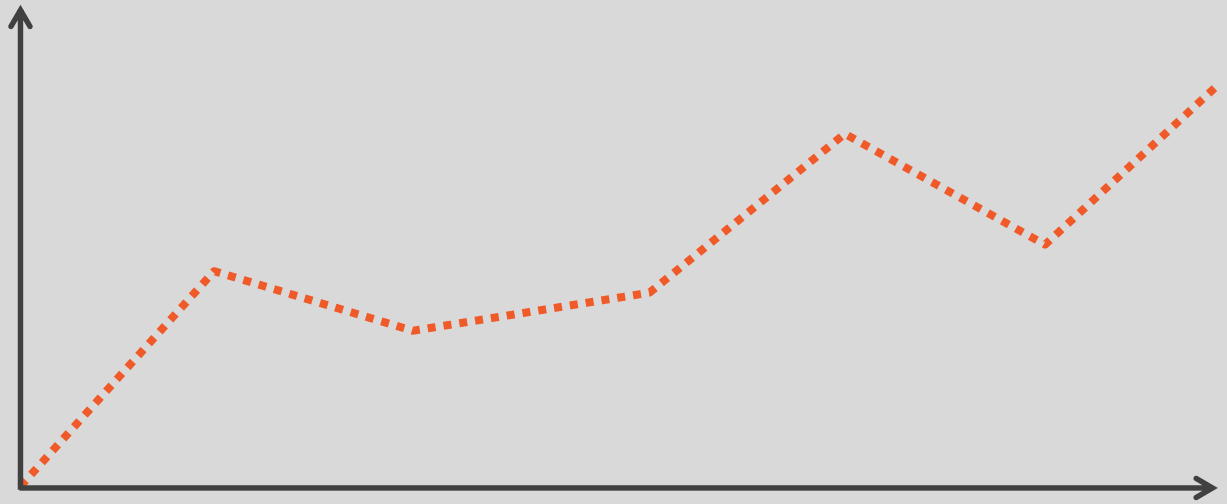


Reactiveness of the Model



$$\beta \approx 0$$

Reactiveness of the Model



$$\beta \approx 1$$

Holt Exponential Smoothing in R

The Main Challenge of Modeling the LFPR

Labor force participation rates cannot cross the 100% mark

A change point in trend is inherent in the data

What Lowers the LFPR?

Disabilities and diseases

Under-education

Unwillingness to participate

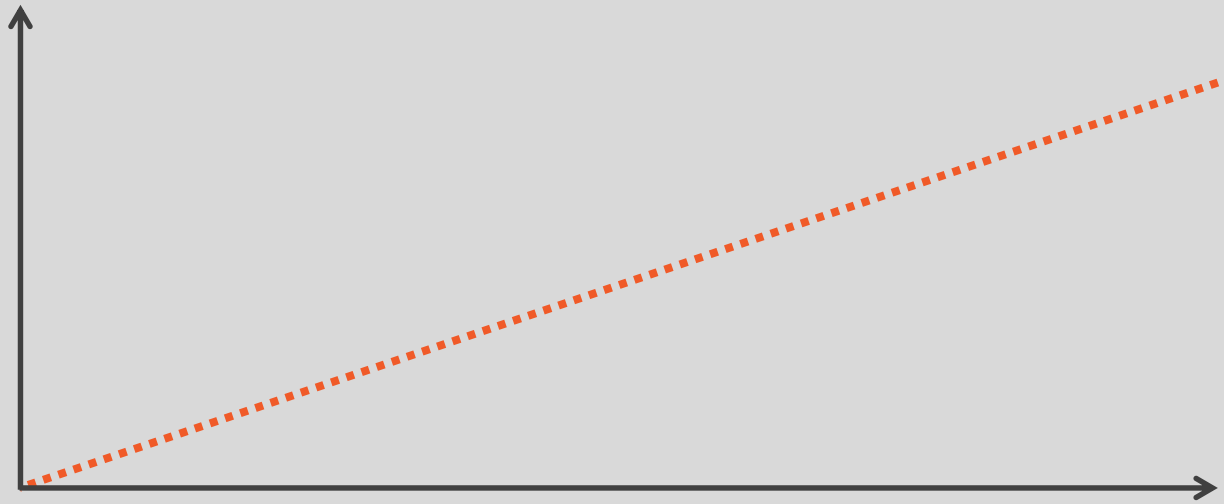
Thresholds are to be determined by experts of the field

- Macro economist or social scientist

Rates cannot grow infinitely

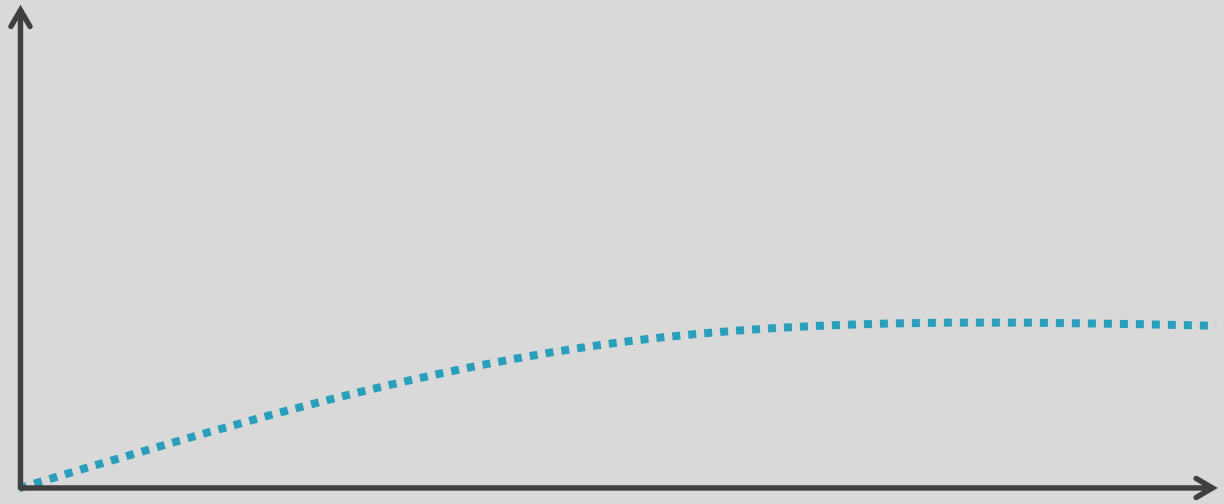
External factors should be incorporated into the model(s)

Damping the Trend in a Holt Model



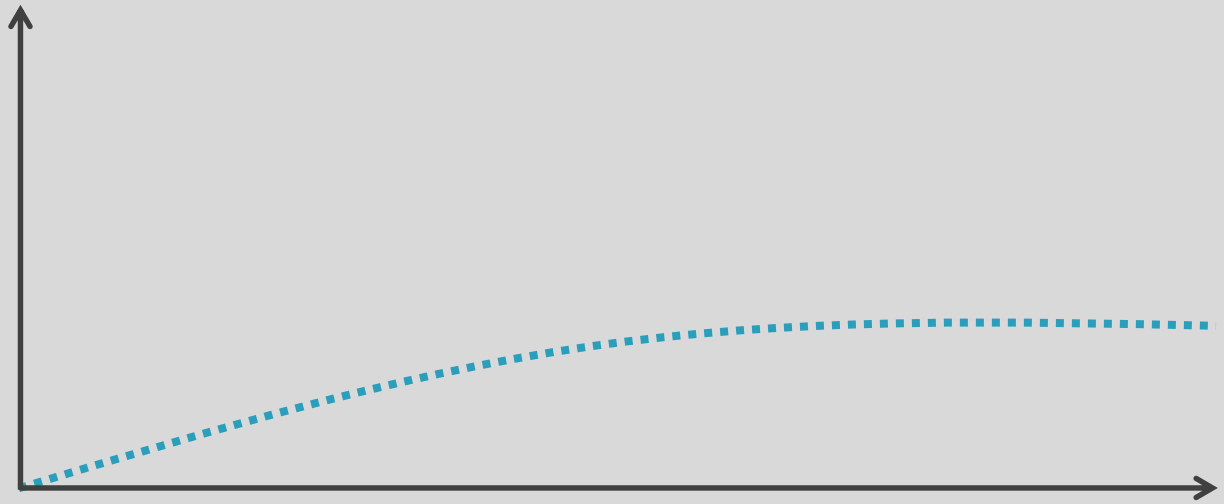
$$\varphi \approx 1$$

Damping the Trend in a Holt Model



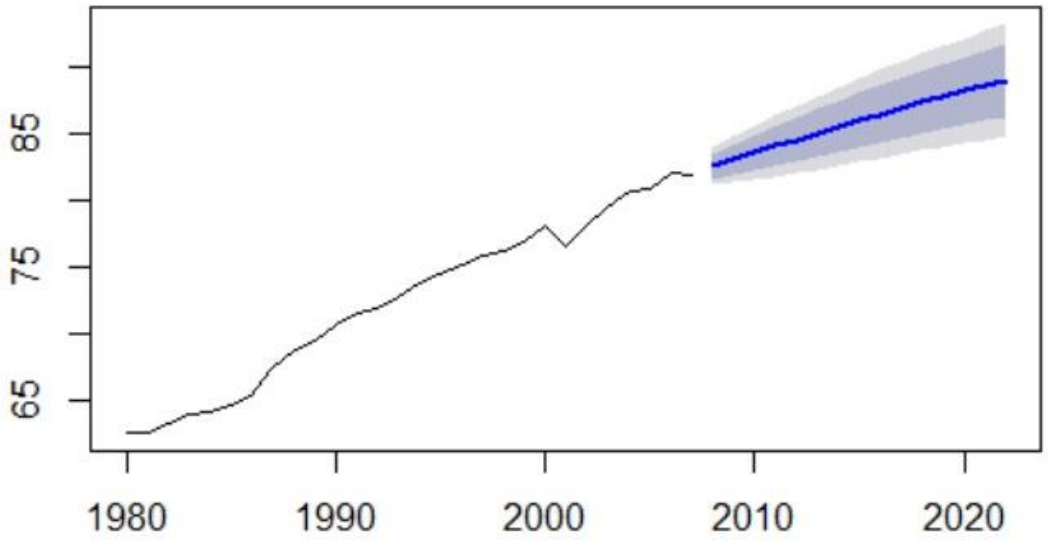
$$\varphi \approx 0$$

Damping the Trend in a Holt Model



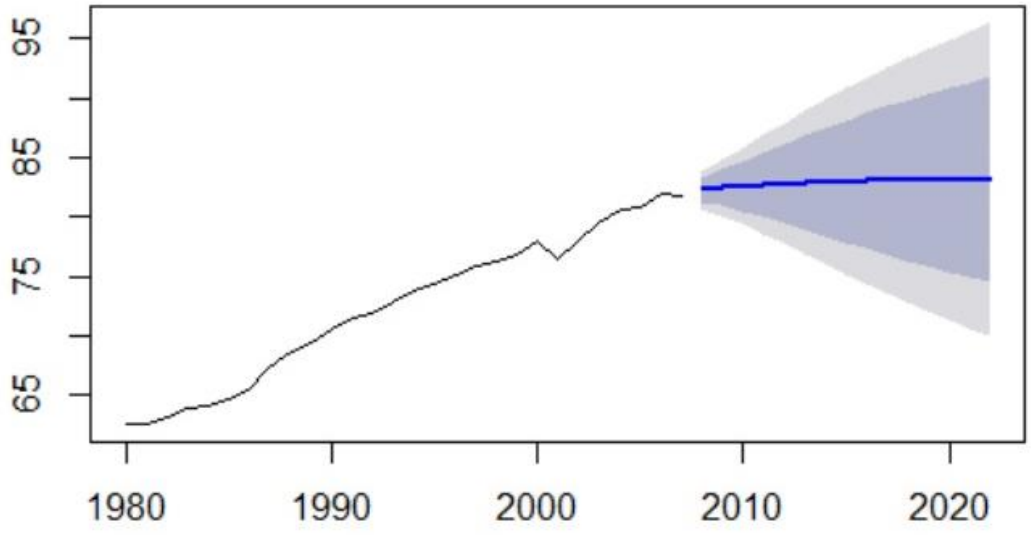
$$0.8 \leq \varphi \leq 0.98$$

Setting the Damping Parameter Value



Estimation of R

$\varphi = 0.979$



Manual set up

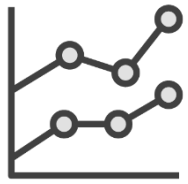
$\varphi = 0.8$

Some scientific disciplines have their standard parameter settings

ARIMA Models for Trending Data

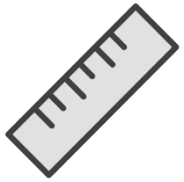
ARIMA Model Parameters

Autoregressive Integrated Moving Average



Autoregression: Captures trend and seasonality (p)

Integration: Captures the differences between the observations (d)



Moving average: Captures movements along a constant mean (q)

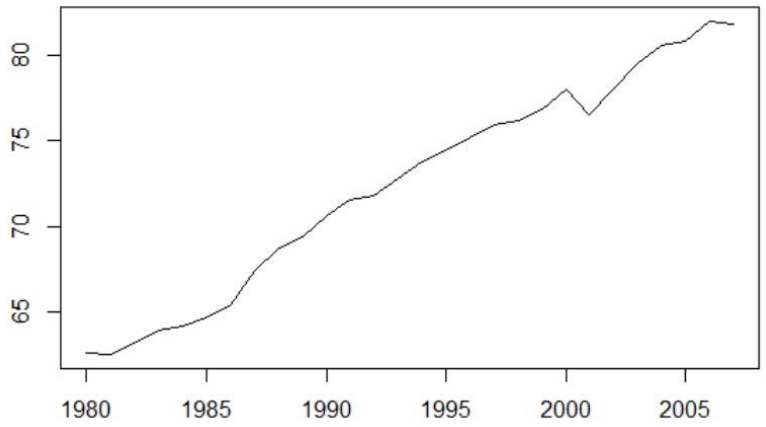
ARIMA Models Are General and Flexible

AR(1) or ARIMA(1,0,0)

Model contains autoregressive component only

MA(1) or ARIMA(0,0,1)

Model contains moving average component only



Modeling the ‘spain’ time series with ARIMA

Trending data

Autocorrelation: Observations at earlier timepoints influence later observations

- Parameter p (AR)

Autocorrelation strongly relates to differencing

- Parameter d (I)

Library ‘forecast’

- `Arima()` – manual
- `auto.arima()` - automated

Differencing

Computing the differences between consecutive observations. As a result of differencing trend and seasonality are eliminated from the time series.

Visualizing Multiple Models

Communicating Data

Assume little statistical knowledge on the audience side

Easy to understand visualizations

Data Visualization with R

`plot()`

Quick and easy plotting solution from R Base

`autoplot()`

Detailed and polished graphs with ggplot2

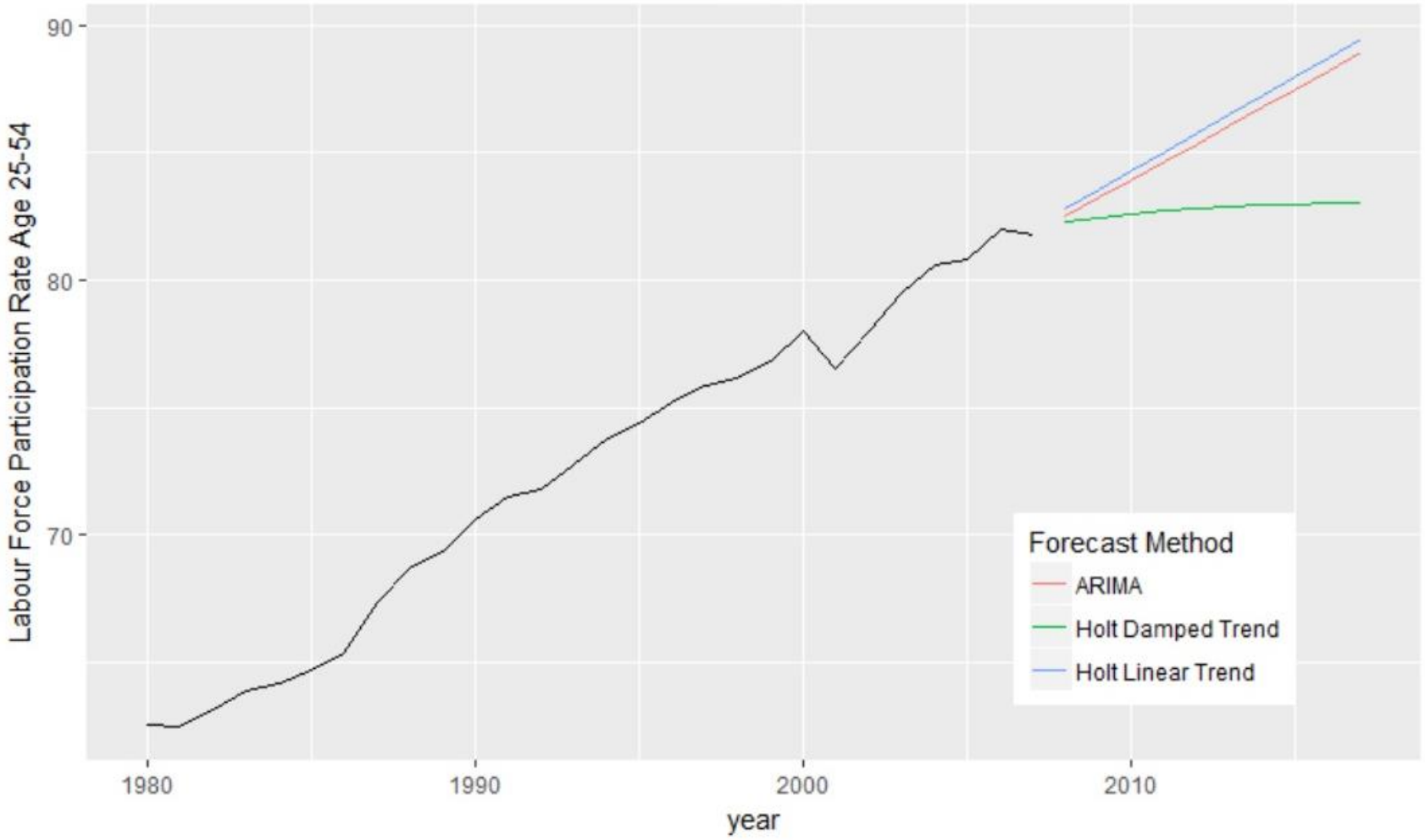
Data Visualizations with ggplot2

ggplot2 is like a sub-language within R

Visualizations are set up layer by layer

The forecast library integrates ggplot2 visualizations

Model Comparison Plot



Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year")+
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Coding a Model Comparison Plot

```
library(ggplot2)

autoplot(spain) +

  forecast::autolayer(holttrend$mean, series = "Holt Linear Trend") +
  forecast::autolayer(holtdamped$mean, series = "Holt Damped Trend") +
  forecast::autolayer(arimafore$mean, series = "ARIMA") +

  xlab("year") +
  ylab("Labour Force Participation Rate Age 25-54") +
  guides(colour = guide_legend(title = "Forecast Method")) +
  theme(legend.position = c(0.8, 0.2)) +
  ggtitle("Spain") +
  theme(plot.title = element_text(family = "Times", hjust = 0.5,
        color = "blue", face = "bold", size = 15))
```

Trending Data

Theoretical background of working with a trending time series

Labor force participation of Spain

- `'holt()'` from `'forecast'`
- `'auto.arima()'` from `'forecast'`

Comparison plot with `'autolayer'` from `'ggplot2'`

Trends come to an end

- Model adjustments with damping parameter phi
- Changepoint