Time Series Modeling II

Agenda

How does ARIMA compare to linear regression?

How you can get a variety of models from just a single series?

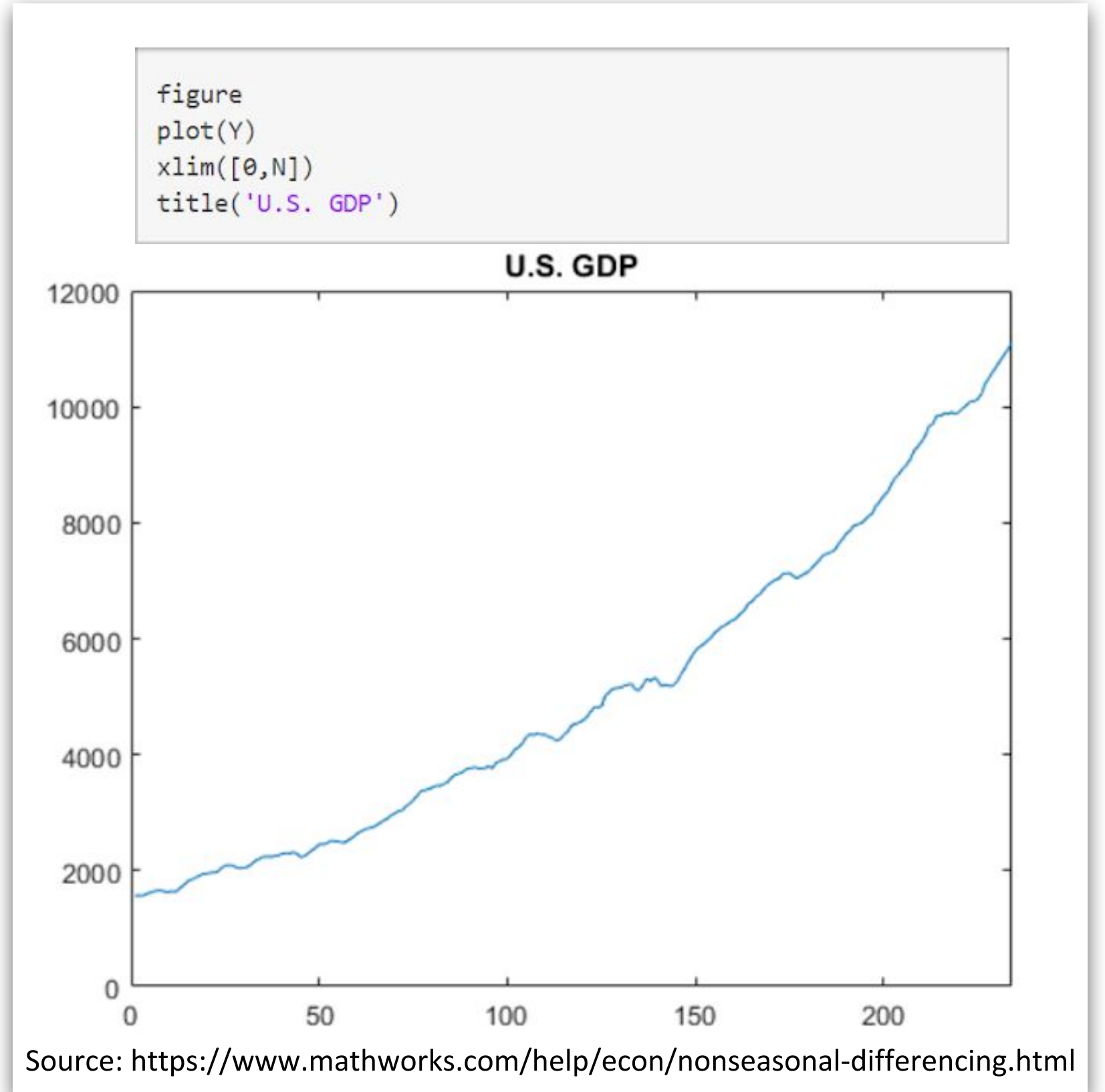How to choose ARIMA parameters for your trading model?

Learning Objectives

- Understand how ARIMA differs from linear regression

- Learn how to apply a variety of models to just a single series

- Understand how to choose the parameters and features of your ARIMA trading model

Concepts in Time Series

What is Stationary data?

"Stationary" means that the statistical structure of the series is independent of time.


```
figure
plot(Y)
xlim([0,N])
title('U.S. GDP')
```

U.S. GDP

Source: https://www.mathworks.com/help/econ/nonseasonal-differencing.html
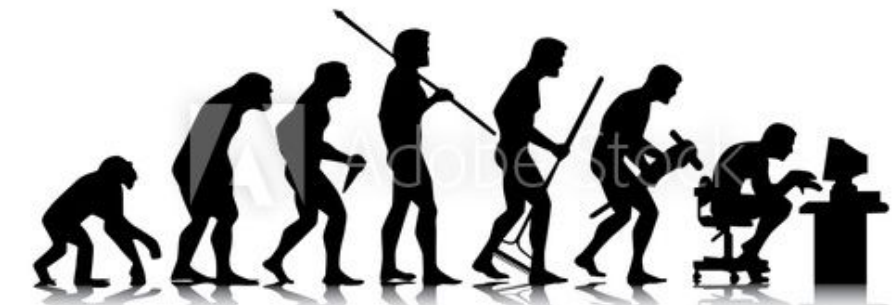
Google Cloud

ARIMA vs Linear Regression

Big Picture

Regression: Model y ~ x

Model y based on the distribution of x

ARIMA: y ~ previous values of y

Model y based on previous values of y

ARIMA vs Linear Regression

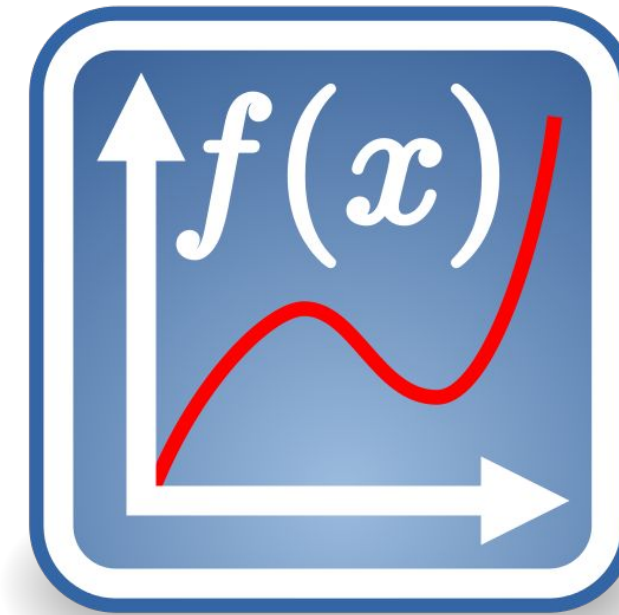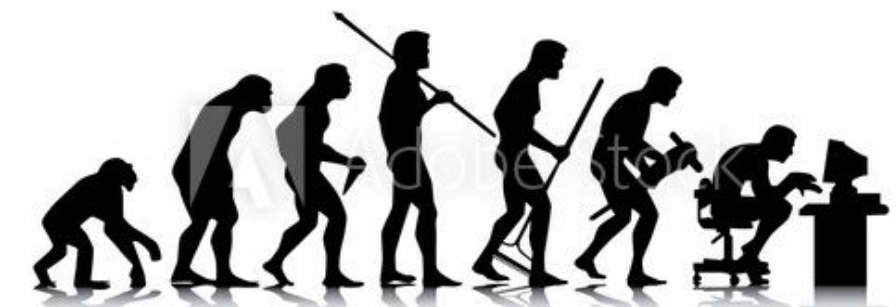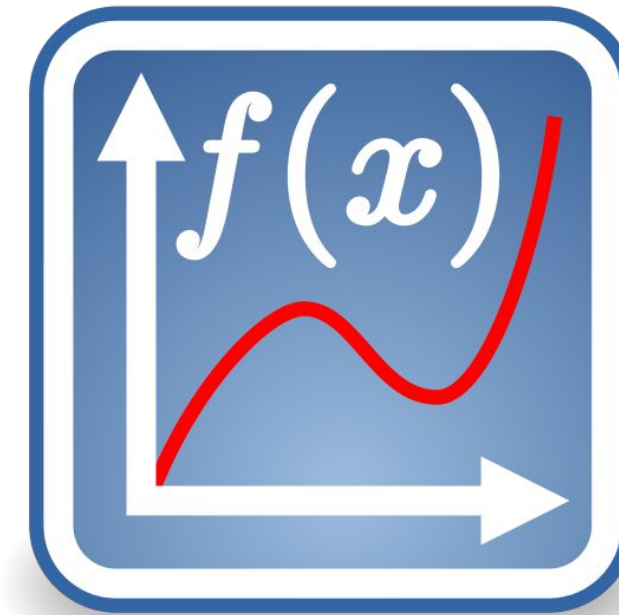**Big Picture**

Regression: Model y ~ x

Model y based on the distribution of x

ARIMA: y ~ previous values of y

Model y based on previous values of y

#17580904

Google Cloud   NEW YORK INSTITUTE OF FINANCE   ESTABLISHED 1922

ARIMA vs Linear Regression

What do they have in common?

Require stationarity

Linear

Correlation between response & dependent variables

Methods to estimate coefficients

Statistical tests to assess the quality

Google Cloud | NEW YORK INSTITUTE OF FINANCE ESTABLISHED 1922

ARIMA vs Linear Regression

What do they have in common?

Require stationarity

Linear

Correlation between response & dependent variables

Methods to estimate coefficients

Statistical tests to assess the quality

Google Cloud   NEW YORK INSTITUTE OF FINANCE   ESTABLISHED 1922

ARIMA vs Linear Regression

What do they do differently?

- There is no natural ordering to the observations in linear regression

- Linear regression uses two different variables

- Linear regression emphasizes one variable depends on the other.

- There is sequential ordering to the observations in time series

- ARIMA uses the same variable, with lagged values of the response

- ARIMA does not need to make this choice

Agenda

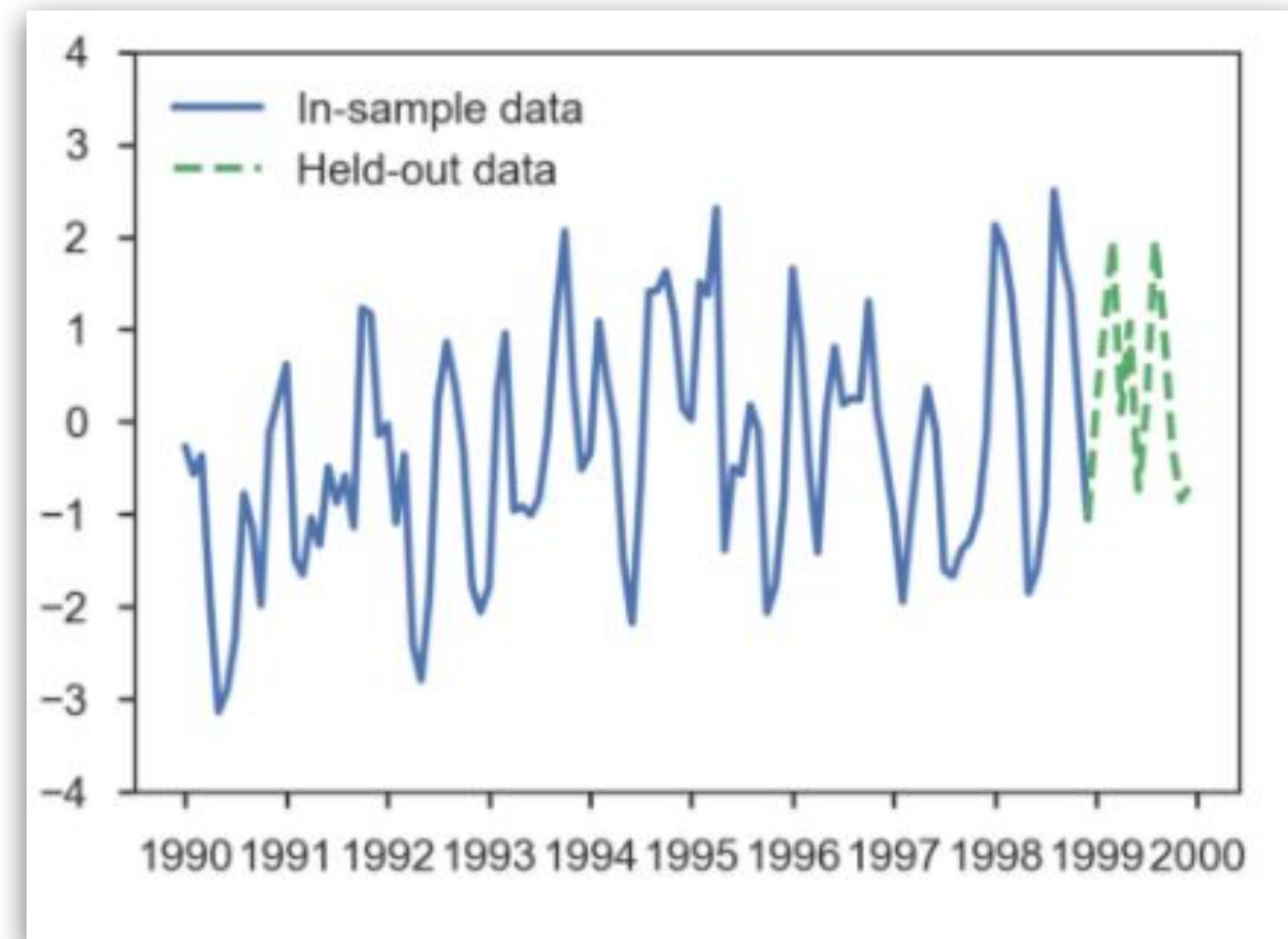How does ARIMA compare to linear regression?

How you can get a variety of models from just a single series?

How to choose ARIMA parameters for your trading model?

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

ARIMA Model

ARIMA (p,d,q) Model

In an ARIMA model, there will be three parameters that will be needed.

ARIMA Model

AR I MA (p,d,q) Model

In an ARIMA model, there will be three parameters that will be needed.

ARIMA Model

AR I MA (p,d,q) Model

In an ARIMA model, there will be three parameters that will be needed.

ARIMA Model

AR I MA (p,d,q) Model

In an ARIMA model, there will be three parameters that will be needed.
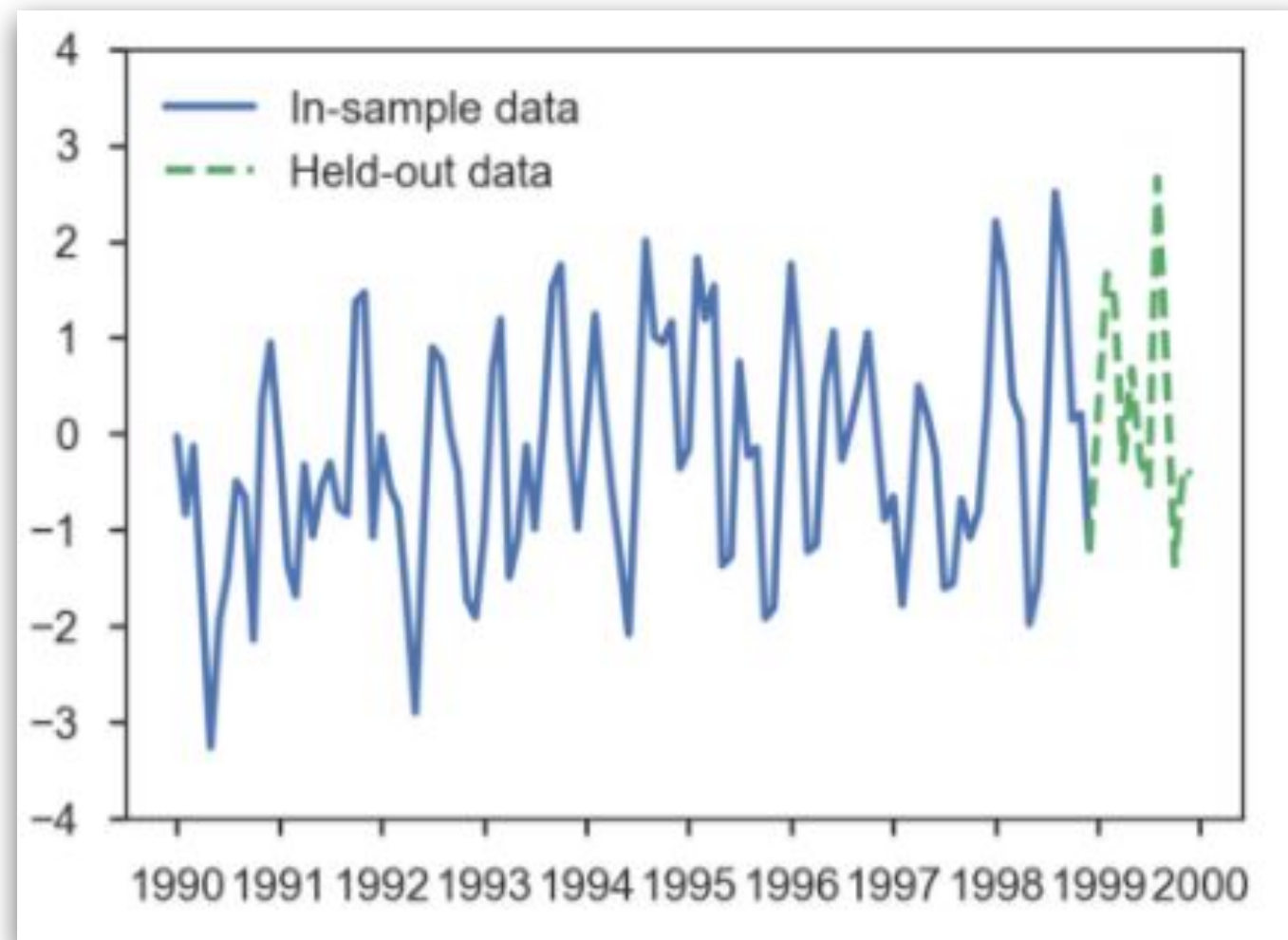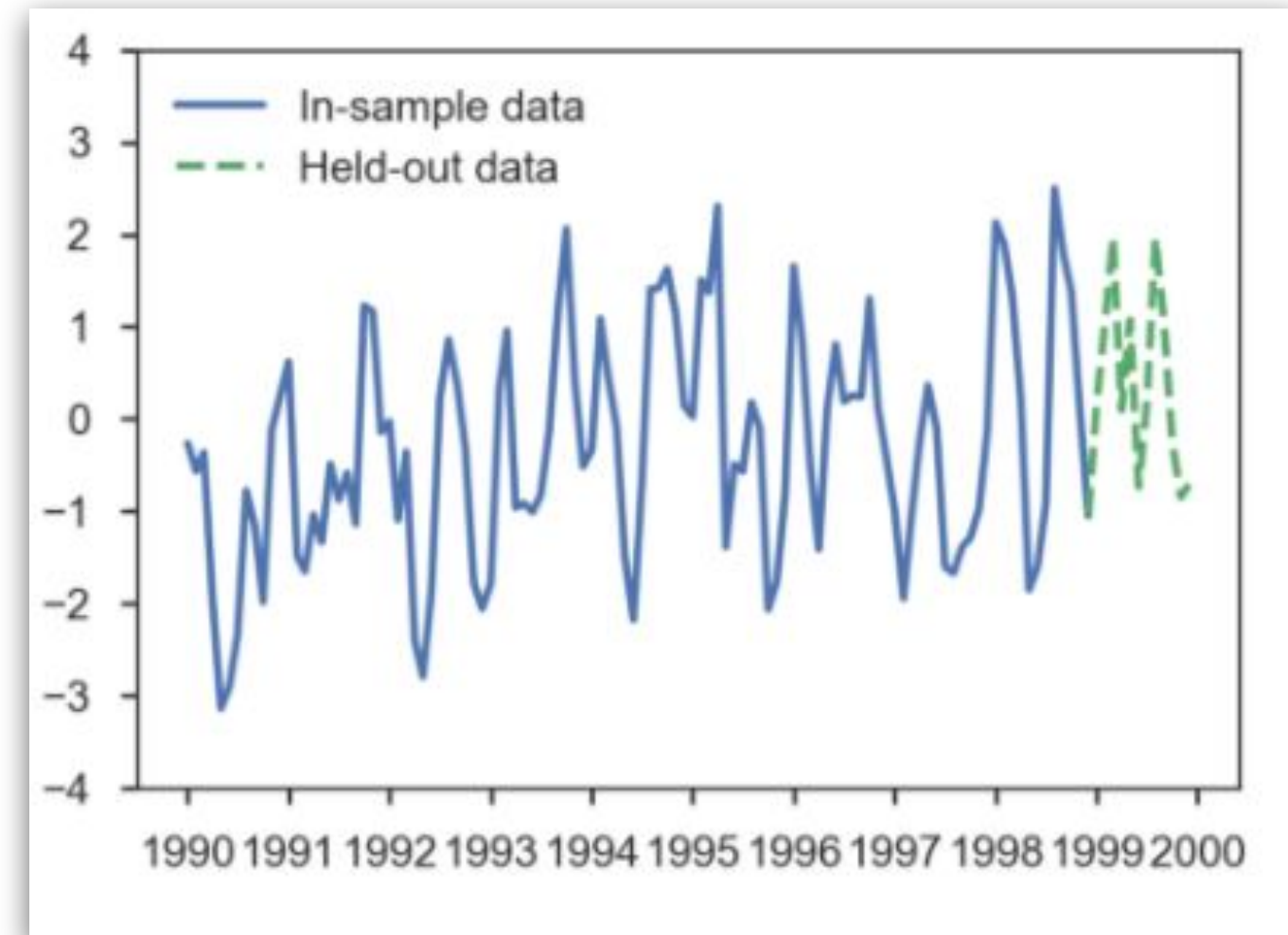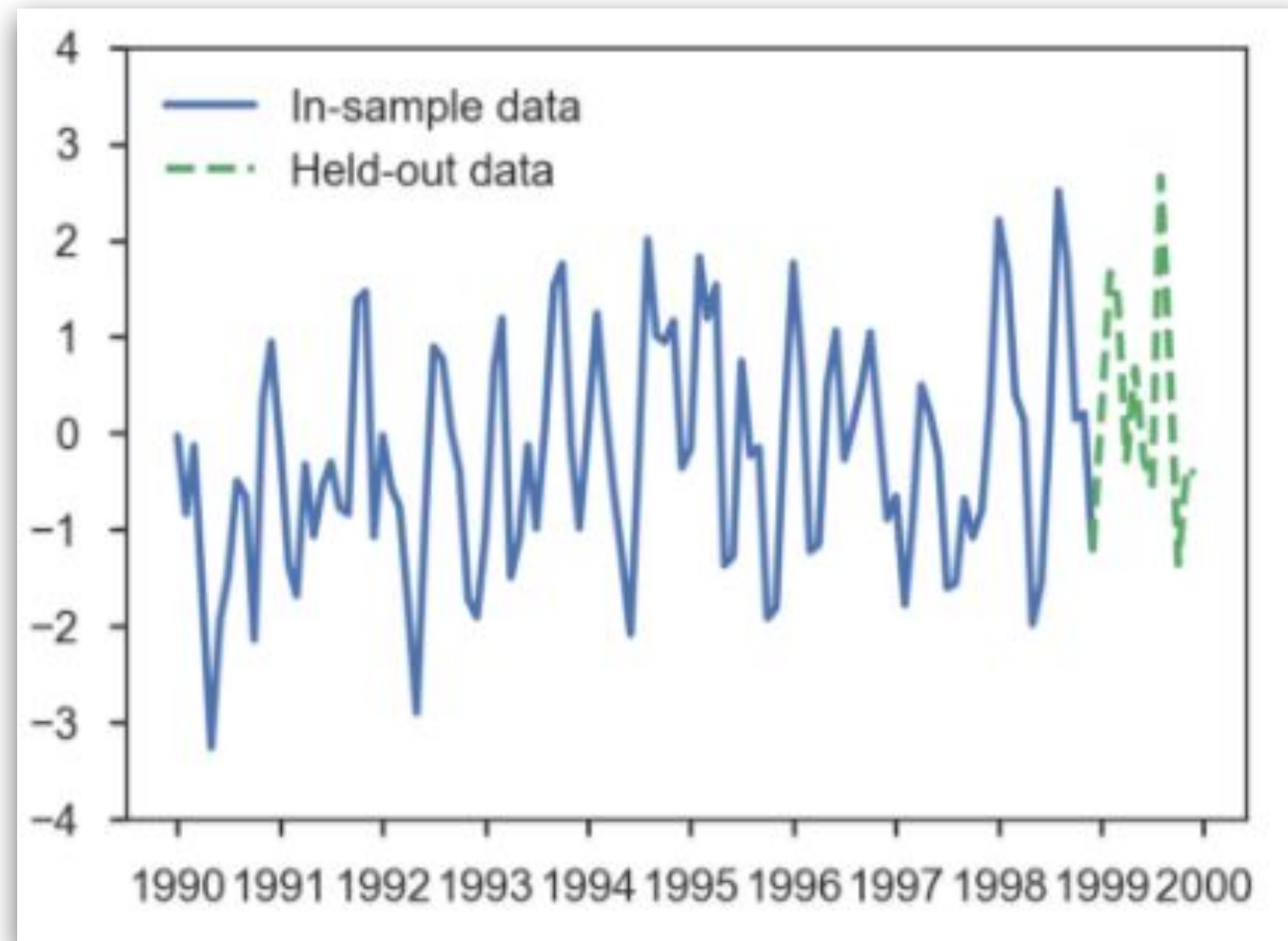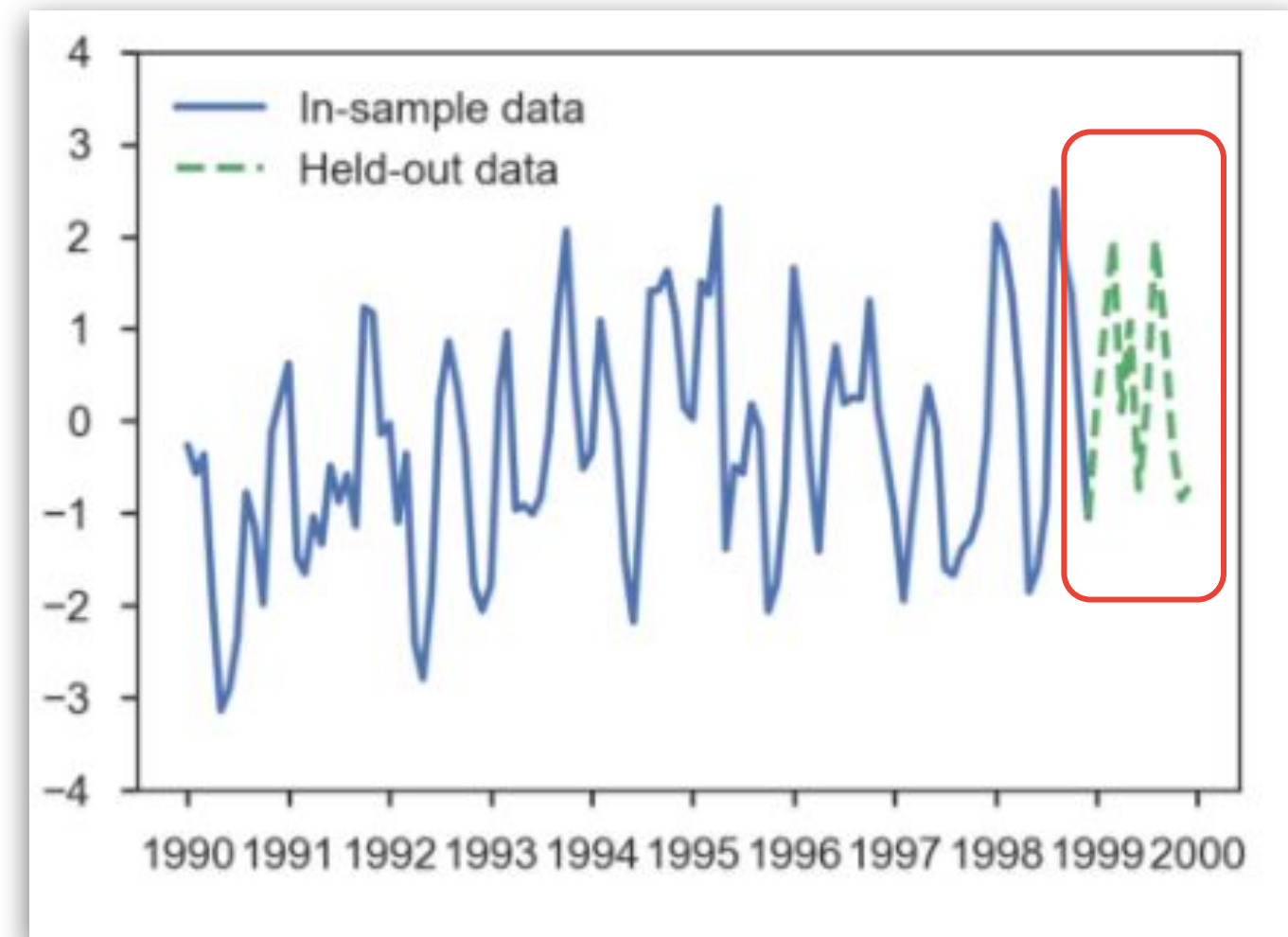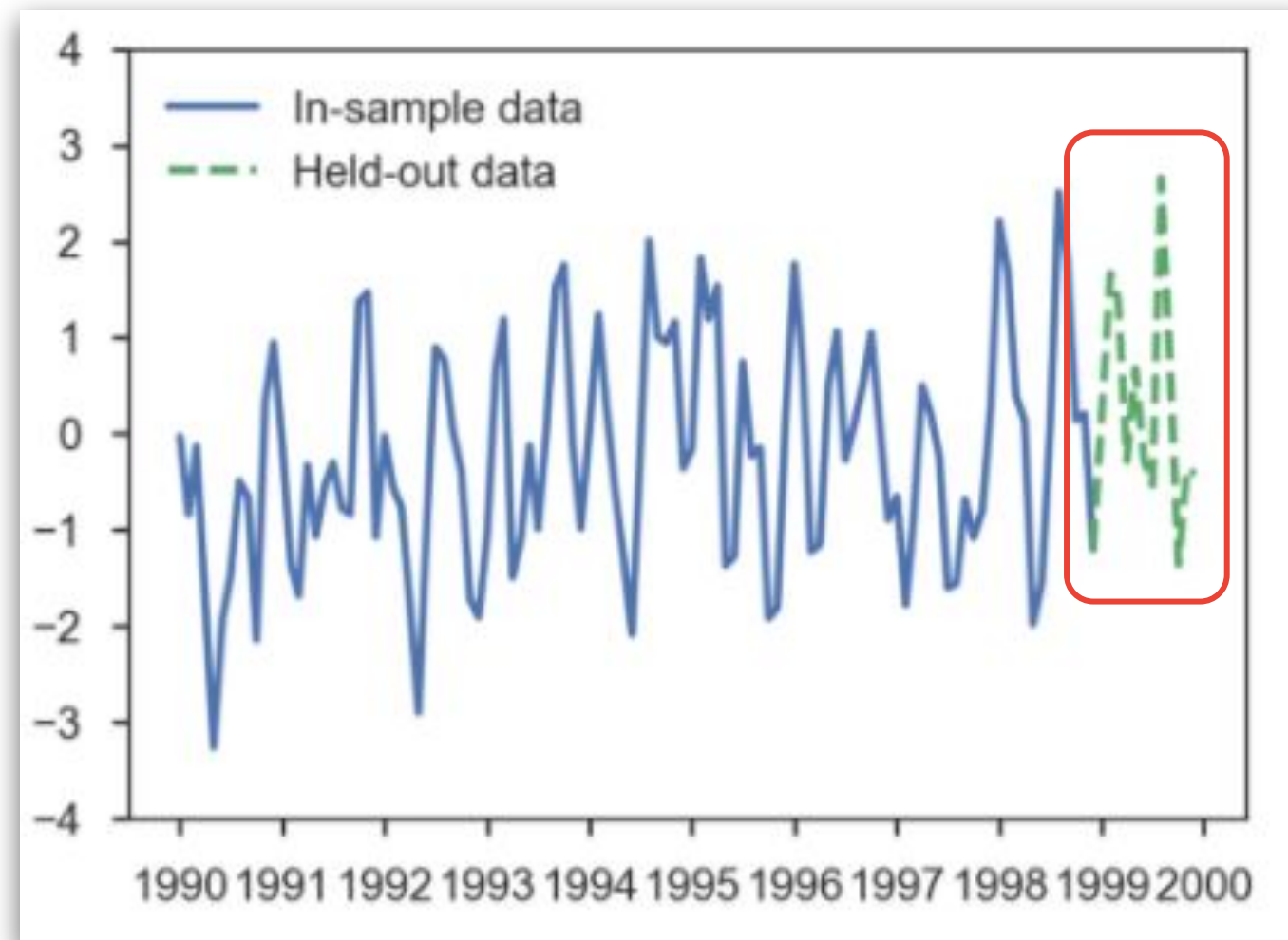
Agenda

How does ARIMA compare to linear regression?

How can you get a variety of models from just a single series?

How to choose ARIMA parameters for your trading model?

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data

- d

- p and q

- the form of ARIMA

- the estimation method

- the best model

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- Most difficult step
- Often have too much data rather than too little
- Need to choose the frequency of data that you need for your model

Google Cloud   NEW YORK INSTITUTE OF FINANCE
ESTABLISHED 1922

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- If d = 0, you are working with levels
  - Prices
  - Yields
- If d =1, you are working with differences
  - Returns and log returns
  - Yield changes

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data

- d

- **p and q**

- the form of ARIMA

- the estimation method

- the best model

- These have to be integers!

- ... but you can choose to set them to zero

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- Variations
  - Fractional estimation
  - Seasonality

Google Cloud   NEW YORK INSTITUTE OF FINANCE   ESTABLISHED 1922

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- **the estimation method** ← - - - - - -
- the best model

- Maximum Likelihood
- Method of moments
- Non-parametric methods
- Simulation

Google Cloud   NEW YORK INSTITUTE OF FINANCE
ESTABLISHED 1922

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- **the best model** ←----------

- May run multiple form of model with different estimation parameters
- May use different software
- Choose model with best quality of fit

Time Series Terminology: Auto Correlation (ARIMA)

Choose:

- the data

- d

- p and q

- the form of ARIMA

- the estimation method

- the best model

Time Series Terminology: Autocorrelation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- Most important
- Tradeoff between timeliness and relevance
- How to divide for training and testing
- Relevance to prediction goal

Google Cloud | NEW YORK INSTITUTE OF FINANCE | ESTABLISHED 1922

Time Series Terminology: Autocorrelation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- Software generally guides this choice
- Can run the model on levels and differences

Google Cloud   NEW YORK INSTITUTE OF FINANCE

Time Series Terminology: Autocorrelation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- Traditional model may oversimplify the analysis
- More complex model may overcomplicate the analysis

Google Cloud

ESTABLISHED 1922
NEW YORK INSTITUTE OF FINANCE

Time Series Terminology: Autocorrelation (ARIMA)

Choose:

- the data

- d

- p and q

- the form of ARIMA

- **the estimation method** ← - - - - - - - - -

- the best model

- Maximum Likelihood
- Method of moments
- Non-parametric methods
- Simulation

Time Series Terminology: Autocorrelation (ARIMA)

Choose:

- the data
- d
- p and q
- the form of ARIMA
- the estimation method
- the best model

- May run multiple form of model with different estimation parameters
- May use different software
- Choose model with best quality of fit

Sensitivity of Trading Strategy

- What is the mean?

- How far does it typically deviate from the mean and how long does it take to get back?

- What is the impact of your choice of p and q?

- How much explanatory power is in the model?

Sensitivity of Trading Strategy

- What is the mean? ←------------

- How far does it typically deviate from the mean and how long does it take to get back?

- What is the impact of your choice of p and q?

- How much explanatory power is in the model?

- Each of your 6 model decisions can affect the mean
- Choice of data can shift mean up or down

Google Cloud   NEW YORK INSTITUTE OF FINANCE ESTABLISHED 1922

Sensitivity of Trading Strategy

- What is the mean?
- How far does it typically deviate from the mean and how long does it take to get back?
- What is the impact of your choice of p and q?
- How much explanatory power is in the model?

- Long-term average, standard deviation and current value tell you where you are in relative terms and also...
- Expected time to revert to mean

Sensitivity of Trading Strategy

- What is the mean?

- How far does it typically deviate from the mean and how long does it take to get back?

- What is the impact of your choice of p and q?

- How much explanatory power is in the model?

- p and q give the form of the model
- High p means market has high memory
- Low p means the past has a low impact on the future

Sensitivity of Trading Strategy

- What is the mean?

- How far does it typically deviate from the mean and how long does it take to get back?

- What is the impact of your choice of p and q?

- How much explanatory power is in the model?

- If p and q are zero your model is a random walk!
- If your model is best fit by AR(1) then there is structure
- Trading strategy is sensitive to model estimates which depend on your choices.

Google Cloud   NEW YORK INSTITUTE OF FINANCE   ESTABLISHED 1922

Lab

Forecasting a stock price
next few days

Lab Objectives

How to import data from GCS

How to setup a Time Series model

How to forecast future using model

How to evaluate results

Google Cloud   NEW YORK INSTITUTE of FINANCE