Quantitative Analysis of Finance I

ECON90033

WEEK 1

COURSE INFORMATION

FINANCIAL ASSET PRICES AND RETURNS

STATISTICAL PROPERTIES OF FINANCIAL DATA

Reference:
HMPY: Ch 1, 2

Dr László Kónya
January 2023

COURSE INFORMATION

__Subject Coordinator and Lecturer:__ *Dr László Kónya*
Office: 460, 4th floor, FBE Building
Phone: 8344 0355
Email: l.konya@unimelb.edu.au

__Consultation hour:__ *THU 14:15 - 15:15 in my office or on Zoom*

__Lectures:__ *There is one two-hour lecture a week.*

FRI 11:00 – 13:00 (The Spot - 4014)

The lectures will be (i) live-streamed and (ii) recorded and made available on LMS, granted that some technical problem does not prevent IT to do so.

If I get sick, I shall make an announcement, record the lecture at home and upload the video on LMS.

Tutorials: *Start in week 1.*

Sign up for one and only one tutorial class by the end of week 1 via the Student Portal.

Subject website: *Time to time, important messages might be uploaded onto the subject LMS website, so visit it regularly.*

The subject guide, the lecture notes, the tutorial materials, and some review resources, can be downloaded from this website in due time.

Although the lecture notes are fairly detailed, they are not meant to substitute for the prescribed and recommended texts.

Software: *R and RStudio.*

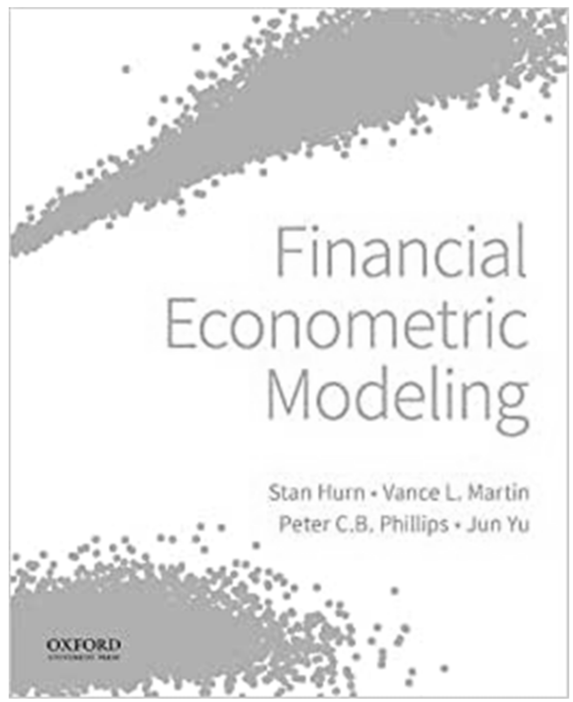
On the exam you will not be tested on the operation of these programs, but (i) you will need to use these programs to complete the assignments, and (ii) some exam questions will be based on R printouts.

Download and install R and RStudio on your computer. You can find the instructions how to do so and how to start using these programs in the "R and RStudio - Part 1 and Part 2" handouts on the subject website.

<u>*Prescribed text:*</u> *Hurn, S., Martin, V.L., Phillips, P.C.B. and Yu, J. (2020): Financial Econometric Modelling, Oxford University Press.*

This book is available in the library. Alternatively, you can rent or purchase it online at

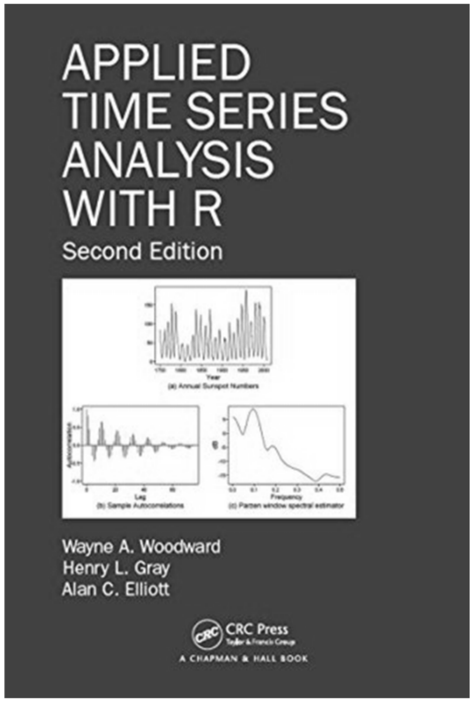
<u>https://www.vitalsource.com/en-au/products/financial-econometric-modeling-stan-hurn-vance-l-martin-v9780190857073?term=9780190857066</u>



Please note that the notations used in the lecture and tutorial materials can be different from the ones used in this text.

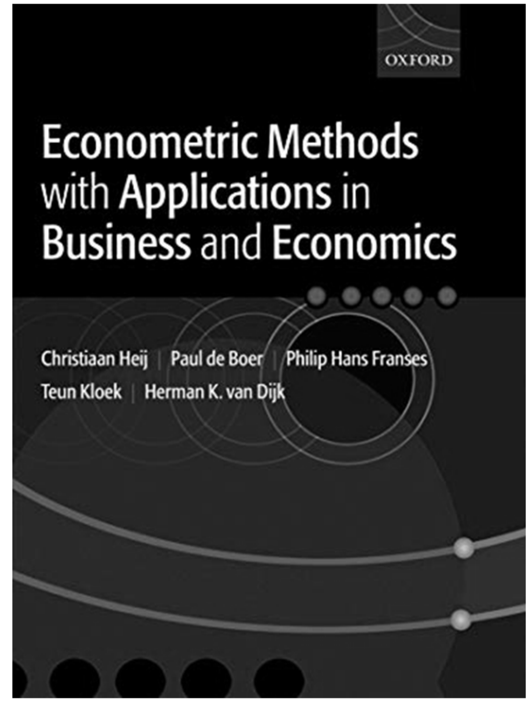
Recommended texts:

Woodward, W.A. et al. (2017)



Background reading:

Heij, C. et al. (2004)



Colonescu, C. (2016):
Principles of Econometrics with R,
self-published,
https://bookdown.org/ccolonescu/RPoE4/RPoE.pdf

Prerequisite:

The official prerequisite for this unit is admission to one of

- *Master of Finance (MC-FINANCE)*
- *Master of Finance – Enhanced (MC-FINANCEH)*
- *Master of Applied Econometrics (MC-AEMTRCS)*
- *Master of Applied Econometrics - Enhanced (MC-AECOENH)*

Students are supposed to be familiar with:

- *Descriptive and inferential business statistics;*
- *Simple linear regression and correlation;*
- *Multiple linear regression (model specification, OLS estimation, hypothesis testing);*
- *Elementary matrix algebra;*
- *Simultaneous equation models;*
- *R programming language.*

Review Appendices A, B and C in Hurn, S. et al. (2020), and the 'Correlation and Regression', Homogeneous Linear Difference Equations' and 'Simultaneous Equation Models' handouts on the subject website.

Subject overview:

Modeling Financial Time Series: weeks 1-5

Week 1: Course Information
Financial Asset Prices and Returns
Statistical Properties of Financial Data

Week 2: Linear Regression Model
Capital Asset Pricing Model

Week 3: Stationarity
Dynamics of Financial Time Series
The Autocorrelation and Partial Autocorrelation Functions
Univariate Time Series Models

Week 4: Deterministic and Stochastic Trends
Spurious Regression
Testing for a Unit Root with the DF τ Tests
Detecting Asset Price Bubbles

Week 5: Introduction to Forecasting

Forecasting with Stationary ARMA Models

Forecasting with Exogenous Predictors

Modeling Risk: weeks 6-9

Week 6: Autoregressive Conditional Heteroskedastic (ARCH) Processes

ARCH and GARCH Models of Conditional Variance

Week 7: Forecasting with GARCH Models

Extensions to the Basic GARCH Model: IGARCH, EGARCH, TGARCH and GARCH-M Models

Week 8: High Frequency Data

Realized Variance

Bipower Variation and Jumps

The Realized GARCH Model

Week 9: Pricing Options

The Black-Scholes Option Price Model

Option Pricing and GARCH Volatility

System Modeling: weeks 10-12

Week 10: Vector Autoregression (VAR)

Granger Causality

Week 11: Impulse Response Analysis

Forecast Error Variance Decomposition

Cointegration

Week 12: Cointegration Testing

Equilibrium Dynamics and Error Correction

Assessment: *Two assignments (1250 and 2250 words for 15% and 25%)* *40%*

Final exam (2 hours) *60%*

Hurdle: *To pass this subject, students must pass the end of semester examination.*

Tutorial classes:

The primary aim of the tutorials is to learn and practice through a wide range of examples how to analyse and forecast time series with R.

Before each tutorial

i. *Attend / watch the previous week's lecture.*
ii. *Go through the relevant tutorial handout. They are self-explanatory and sufficiently detailed, so follow the instructions, and reproduce the illustrative example(s).*
iii. *If you need help, ask your tutor before or during the tutorial class for assistance, but do not expect your tutor to cover the entire handout.*

Assignments:

There will be two assignments for 15% and 25% credit, respectively.

i. *Online submission via Canvas.*
ii. *Students can work alone or in a group of two (not three or four …).*
iii. *Students in a group must submit a single copy of their assignment and will get the same assignment marks.*
iv. *No late assignments are accepted, and no extensions will be given.*

Final exam at the end of the semester:

It is worth 60% of the final grade for this subject.

i. *It will be a 2-hour exam during the University's normal end of semester assessment period. The exact date, time and location of the exam will be provided by the University's administration later in the semester.*
ii. *The exam will cover all materials discussed during lectures and tutorials throughout the semester. There will be no surprises, the exam questions and tasks will be similar in terms of style and difficulty to those in the tutorial problem sets and in the assignments.*

iii. It will be an open-book exam, so formula sheet and statistical tables will not be provided on the exam.

iv. On the exam students will neither be asked nor tested on how to use R / RStudio, but they will need to be familiar with R printouts.

v. Students must pass the exam, i.e., to achieve 50% of the total exam mark, to successfully complete the subject.

vi. Supplementary exam will not be provided in case of absence during the examination period, unless it is due to serious illness or some other legitimate reason. In those exceptional cases apply for special consideration.

(https://students.unimelb.edu.au/admin/special-consideration).

INTRODUCTION TO QUANTITATIVE ANALYSIS OF FINANCE

- In terms of data collection, we distinguish three main types of data.

  Cross-sectional data: observations on some variable (or variables) of interest measured across a sample of individuals, households, firms, cities, countries etc. at the same point in time.

  E.g., the population of Australian states and territories at the end of June 2018.

  Time-series data: observations on some variable (or variables) of interest collected over discrete and usually regular intervals of time (every day, week, month, quarter, year etc.).

  E.g., estimates of mid-year population of Australia 1960-2022.

  Panel data (or longitudinal data): the combination of cross-sectional and time-series data, i.e., a given set of cross-sectional units are observed repeatedly at multiple points in time.

  E.g., population of CANZUK countries (Canada, Australia, New Zealand and UK) 1960-2022.

- Most financial data are time series data, so in this course we mainly apply time series econometrics.

  According to classical time-series analysis an observed time series is the combination of some *pattern* and *random variations*,

  and the aim is to separate them from each other in order to

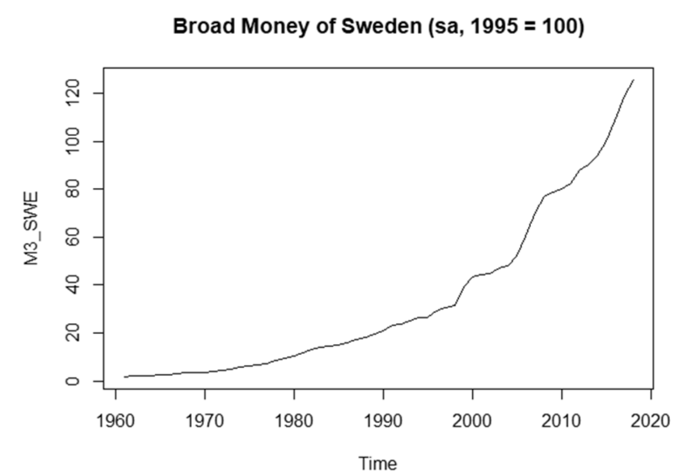
  (a) describe the historical pattern in the data and
  (b) to prepare forecasts by projecting the revealed historical pattern into the future.

- The pattern itself is likely to contain some or all of the following three components: *trend*, *seasonal* and *cyclical*.

Trend:

It is the long-term general change in the level of the data.

It can be linear ($y_t = a + bt$), but it does not have to be, and it might change direction.



Seasonal variations:

Regular wavelike fluctuations of constant length, repeating themselves within a period of no longer than a year.

They are usually associated with the four seasons of the year, but they may also refer to any systematic pattern that occurs during a month, a week or even a single day.

There are two types of seasonality: additive and multiplicative. In the former case the amplitude of the seasonal variation is independent of the level, whereas in the latter case it is proportional to the level.

$$y_t = T_t \times S_t \times \ldots$$

$$\downarrow$$

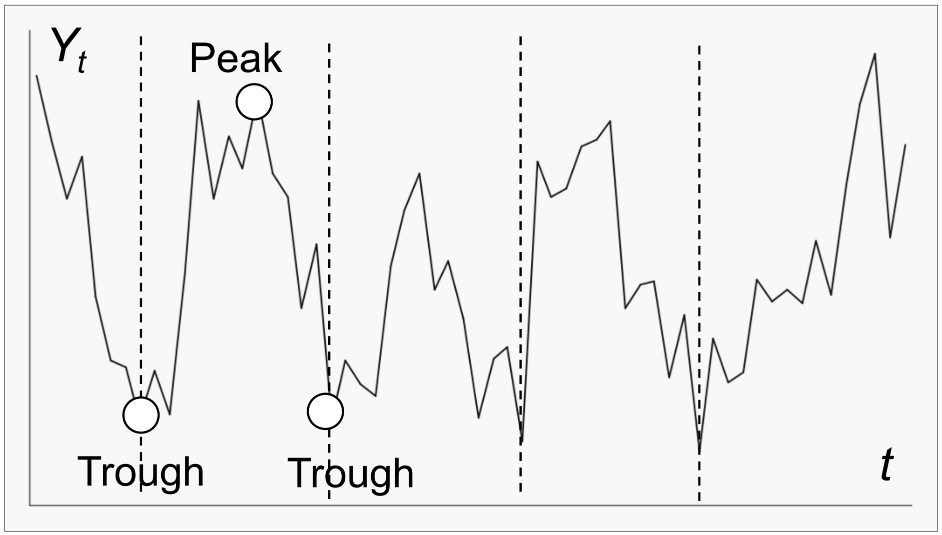
$$\ln(y_t) = \ln(T_t) + \ln(S_t) + \ldots$$

Multiplicative: the fluctuations around the trend tend to increase with the trend.

Additive: the fluctuations around the trend appear to have the same magnitude.

Cyclical variations:

Wavelike movements, quasi regular fluctuations around the long-term trend, lasting longer than a year. They are often attributed to business cycles, i.e., to ups and downs in the general level of business activity.



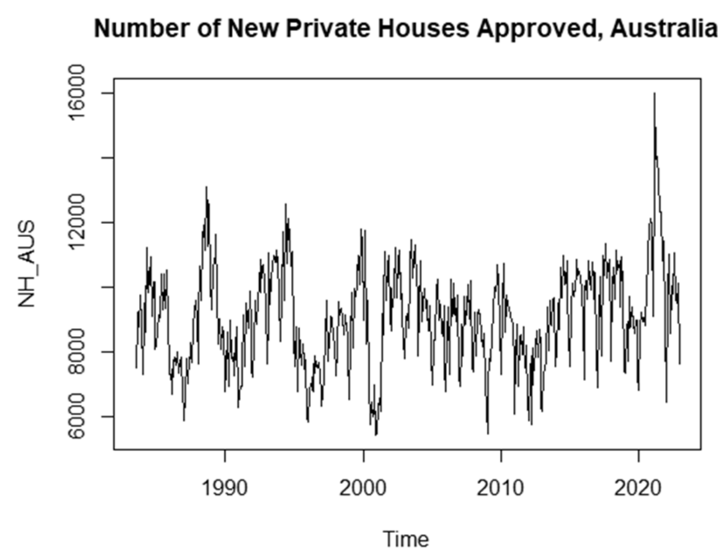
The time gap between consecutive troughs is the *length* of the cycle,

while the vertical distance between the through and the peak is the *amplitude* of the cycle.

The length and the amplitude can both change from cycle to cycle.

Expansion phase: trough → peak

Recession phase: peak → trough



Note: Seasonal and cyclical variations might look very similar in their appearance. However, while seasonal variations are perfectly regular and occur over periods no longer than a year, cyclical variations change in their intensity (amplitude) and/or duration and last longer than a year. For this reason, it is far more difficult to study and predict the cyclical component than the seasonal component.

- The four components of a time series (*T*: trend, *S*: seasonal, *C*: cyclical, *R*: random) can be combined in different ways. Accordingly, univariate time series models used to describe the observed data (*Y*) can be

Additive:

$$y_t = T_t + S_t + C_t + R_t$$

(or some combination of the two)

Multiplicative:

$$y_t = T_t \times S_t \times C_t \times R_t$$

For example, if the trend is linear, these two models are:

$$y_t = (a + bt) + S_t + C_t + R_t$$

S, *C* and *R* are *absolute* deviations from the trend, so they do not depend on the level of the trend.

$$y_t = (a + bt) \times S_t \times C_t \times R_t$$

S, *C* and *R* are *relative* deviations from the trend, so the higher the trend the larger they are.

Traditionally, there are two types of methods for identifying the pattern,

Smoothing:

The random fluctuations are removed from the data by smoothing the time series.

Decomposition:

The time series is broken into its components and the pattern is the combination of the systematic parts.

- Classical time series analysis is purely descriptive and confines all stochastic movements to the random component (R).

  ⟷ The modern approach to time series analysis assumes that an observed time series has been generated by a sequence of random variables $\{Y_t\}$, called stochastic process.

  → Each observation (y_t , $t = 1, 2, \ldots, T$) is supposed to be drawn from the probability distribution corresponding of the corresponding Y_t random variable.

  → The actual time series $\{y_t\}$, is a particular realization of the underlying stochastic process, which is supposed to have begun in the infinite past and to continue forever.

- In this subject we heavily rely on financial econometrics.

Financial econometrics is an interdisciplinary subject, the combination of finance, economics, applied mathematics, statistics and time-series econometrics.

Its aim is to apply various, traditional and modern statistical methods to financial market data.

FINANCIAL ASSET PRICES AND RETURNS

- The origins of financial econometrics can be traced back to early empirical studies of asset prices, stock prices, bond yields, and interest rates.

  An asset sale is the purchase of individual assets and liabilities (owned and owed cash, stocks, bonds, mutual funds, bank deposits) and asset prices are crucial in finance as they represent the cost of a financial instrument (monetary contract).

  Since any change in the price of a financial instrument results in a profit or loss for an investor, it is important to set up econometric models in order to predict future price movements and to study how the price of an asset relates to the prices of other assets (portfolio risk management).

Ex 1:

Consider the quoted prices for the stocks of the Australian multinational iron ore company, BHP Group Limited, obtained from Yahoo Finance on 3 May 2023 (*https://au.finance.yahoo.com/quote/BHP.AX?p=BHP.AX*).

Under the current price, several other prices and summary measures are reported:

Previous close	43.63	Market cap	309.634B
Open	43.40	Beta (5Y monthly)	0.82
Bid	43.17 x 68400	PE ratio (TTM)	7.94
Ask	43.17 x 105900	EPS (TTM)	5.44
Day's range	42.96 - 43.42	Earnings date	20 Feb 2023
52-week range	35.83 - 50.21	Forward dividend & yield	3.92 (8.86%)
Volume	1,540,113	Ex-dividend date	09 Mar 2023
Avg. volume	8,865,630	1y target est	47.11

Previous day *closing* price and the *opening* price of the stock on the given day.

Bid is the highest price a potential buyer is willing to pay for the given quantity of stocks,

and *ask* is the lowest price a stockholder is willing to sell the given quantity of stocks for.

Day's range and *52-week range* are the price range on the given day and the previous 52 weeks, respectively.

Volume and *Avg. volume* are the number of shares traded on the given day and on average, respectively.

Market cap is the number of shares outstanding multiplied by the current share price of a stock.

Ex-dividend date is the day on which shares of a given stock no longer come with the right to collect the next dividend.

1y target est. is the price that analysts have predicted the stock will be one year from now.

Beta (5Y monthly) is the 5-year volatility of the stock relative to the volatility of the market.

PE ratio (TTM) is the price-to-earnings ratio of the stock in the trailing twelve months.

EPS (TTM) is the earnings per share ratio of the stock in the trailing twelve months.

Earnings date is the the next earnings release date.

Forward dividend & yield are annual estimate of the stock price for the next year in AUD and relative to the stock price, respectively.

- Aggregate summary measures of the performances of stock markets are known as stock market indices.

Each stock market index measures the combined value of many stocks traded on the given stock market in terms of some weighted average price of these stocks.

Based on the weights, there are two types of indices.

Price-weighted indices:

The weight assigned to each stock is proportional to its share price.

For example,

Dow Jones Industrial Average (*DJIA*)

Nikkei 225 Index (*NKX*)

Value-weighted indices:

The weight assigned to each stock is proportional to the total market value of its outstanding equity.

For example,

Financial Times Stock Exchange 100 Index (*FTSE*)

Hang Seng Index (*HSX*)

Standard and Poor's Composite 500 (*S&P 500*)

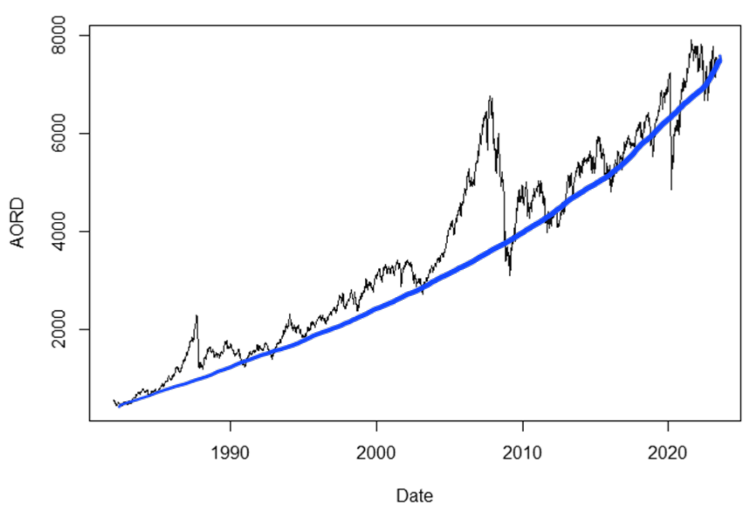
Ex 2:

The ASX All Ordinaries (*AORD*) Index is the market capitalisation (value) weighted average price of the shares of the 500 largest companies listed on the Australian Securities Exchange (ASX), and it is the main summary measure of the movements of share values that result when shares are traded on ASX.

Weekly *AORD* (*P* in AUD) from 1982 week 1 to 2023 week 17 has been downloaded from *https://au.investing.com/indices/all-ordinaries-historical-data*.

a) Plot *P* and describe its data pattern.

```
P = ts(Price, frequency = 52, start = c(1982, 1))
plot.ts(P, xlab = "Date", ylab = "AORD",
        col = "blue")
```



Out of the three possible components of data patterns, the trend component is the most visible this time.

This upward trend clearly does not follow a straight line, i.e., it is not linear, but rather exponential,

$$P_t = P_0(1+R)^t$$

where *R* is the rate of exponential growth.

This constant growth rate can be estimated from the P_0 and P_t:

$$P_t = P_0(1+R)^t \longrightarrow \hat{R} = \sqrt[t]{\frac{P_t}{P_0}} - 1 = \sqrt[2156]{\frac{7538.4}{557.8}} - 1 = 0.001208$$

P_t = final price

P_0 = first price

(2156 is the number of weeks during the sample period)

→ The average weekly return of P is 0.1208%, which implies

$$\hat{R}(52.14) = (1+\hat{R})^{52.14} - 1 = 1.001208^{52.14} - 1 = 0.064971$$

(365 / 7 = 52.14)

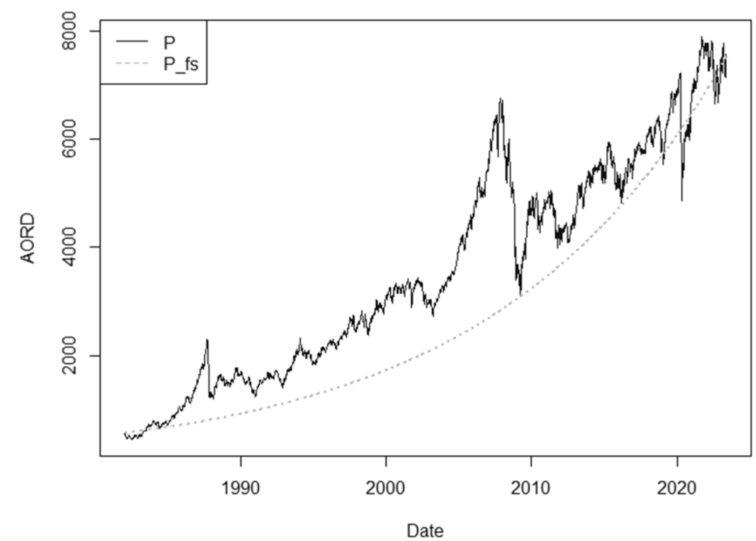
an annualized simple return of 6.4971%.

P_0

Using R-hat and the initial value of P, project P for every week of the sample period.

```
t = 1:length(P)
R = (P[length(P)]/P[1])^(1/length(P)) – 1
P_fs = ts(P[1]*(1+R)^t, frequency = 52, start = c(1982, 1))

plot.ts(P, xlab = "Date", ylab = "AORD", col = "blue")
lines(P_fs, type="l", lty = 3, lwd = 2, col = "green")
legend("topleft", legend = c("P", "P_fs"),
        col = c("blue", "green"), lty = c(1:3))
```



Since *P_f* is based exclusively on P_{first} and P_{last}, *P* and *P_fs* are equal in the first and last weeks of the sample period. Otherwise, $P - P\_fs$ tends to be positive, i.e., *P_fs* systematically underestimates *P.*

- In finance, dollar return is the change in price of an asset, investment, or project between time $t - k$ and time t,

  $\boxed{P_t - P_{t-k} \quad (k > 0)}$ It can be positive (profit), negative (loss), or zero.

  It has two shortcomings, which potentially prevent comparisons across assets, financial markets, and time periods. Namely,

  i. Prices depend on the units of measurement.
  ii. The dollar return also depends on k (holding period).

  A scale-free alternative to the dollar return is the simple (net) return, which is the proportional (percentage) change in price of an asset,

$$\boxed{R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1} \longrightarrow \boxed{1 + R_t = \frac{P_t}{P_{t-1}}}$$

This is the simple gross return,

i.e., the value at time t of \$1 invested at $t - 1$.

- The multi-period return between time $t-k$ and time t is

$$R_t(k) = \frac{P_t}{P_{t-k}} - 1 = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \ldots \times \frac{P_{t-k+1}}{P_{t-k}} - 1$$

$$= (1+R_t)\times(1+R_{t-1})\times\ldots\times(1+R_{t-k+1}) - 1 = \prod_{j=0}^{k-1}(1+R_{t-j}) - 1$$

which is the product of the intermediate one-period gross returns minus one.

For example, if the original data frequency is monthly, then the simple return for a holding period of 1 year, called annualized simple return, is

$$R_t(12) = \prod_{j=0}^{11}(1+R_{t-j}) - 1 = (1+R_t)^{12} - 1$$

granted that the monthly simple return (R_t) has not changed for 12 months.

From the multi-period return formula,

$$\ln\left(1+R_t(k)\right) = \sum_{j=0}^{k-1}\ln(1+R_{t-j})$$

and the terms in this summation are logarithmic simple returns.

- The logarithmic returns are called log-returns.

→ The one-period (gross) log-return on an asset at time *t* is

to get a multi period return?

$$\boxed{r_t = \ln(1+R_t)} = \ln P_t - \ln P_{t-1} = \Delta \ln P_t$$

→ take the period you're interested
→ log the price
→ subtract the logged price of the asset from k periods prior

while the *k*-period log-return between time $t-k$ and time *t* is

$$\begin{aligned} r_t(k) &= \ln P_t - \ln P_{t-k} \\ &= (\ln P_t - \ln P_{t-1}) + (\ln P_{t-1} - \ln P_{t-2}) + \ldots + (\ln P_{t-k+1} - \ln P_{t-k}) \\ &= \sum_{i=0}^{k-1} r_{t-i} \end{aligned}$$

which is the sum of the intermediate one-period log-returns.

The *k*-period log-return is actually a *k*-period compound return, so the logarithmic returns are also called continuously compounded returns.

If the log-return is constant (*r*), the *k*-period log-return is

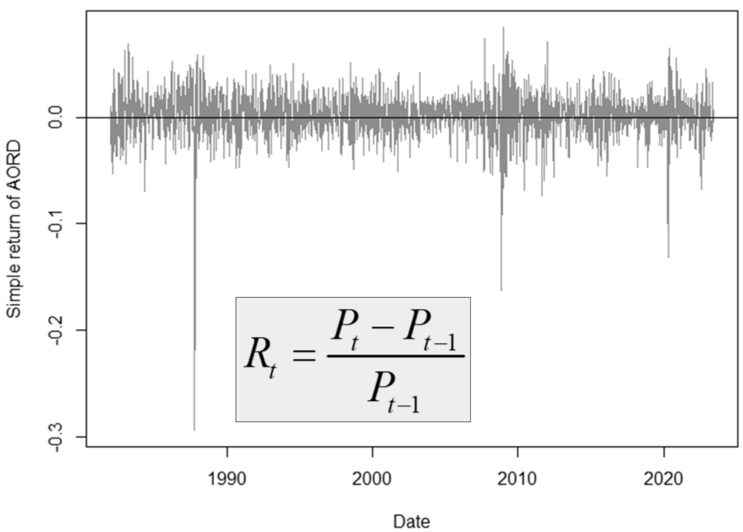
$r_t(k) = kr$ → The annualized log-return is $r_t(365) = 365r$

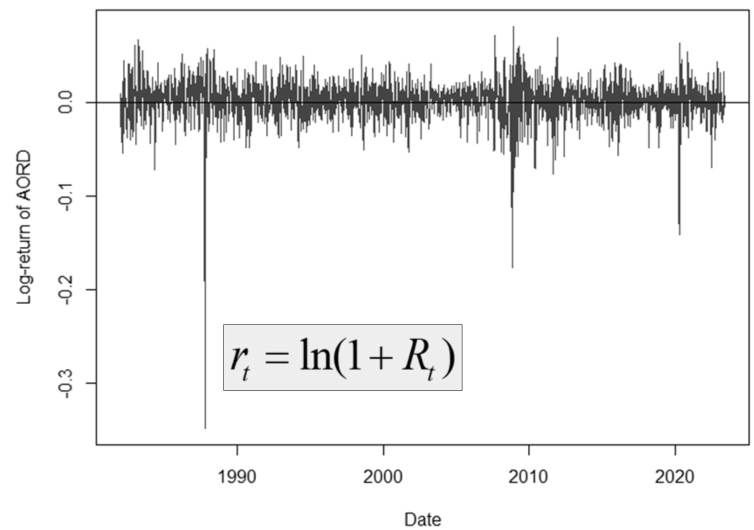
(Ex 2 cont.)

b) Calculate and plot the one-period simple returns and log-returns of the All Ordinaries Index.

```
R = P/lag(P, -1) – 1
plot.ts(R, xlab = "Date", ylab = "Simple return
        of AORD", col = "coral1")
abline(h = 0)
```

```
r = ts(log(1+R), frequency = 52,
        start = c(1982, 1))
plot.ts(R, xlab = "Date", ylab = “Log-return
        of AORD", col = "coral1")
abline(h = 0)
```





The two plots look identical because for small *x*,

$$\boxed{\ln\left(1+x\right) \approx x}$$

- Suppose that the simple return on some asset is constant, *R*,

$$R = \frac{P_t}{P_{t-1}} - 1 \longrightarrow P_t = P_0(1+R)^t$$

In order to make further manipulations easier, we change the base of the P_t exponential function from (1 + *R*) to *e* (Euler number ≈ 2.71828) and use the natural exponential function,

$$P_t = P_0(1+R)^t = P_0 e^{rt} \longrightarrow r = \ln(1+R) = \ln P_t - \ln P_{t-1}$$

Constant log-return

$$\downarrow$$

$$\ln P_t = \ln P_0 + t\ln(1+R) = \ln P_0 + tr \longrightarrow r = \frac{\ln P_t - \ln P_0}{t}$$

(Ex 2 cont.)

c) The constant log-return of the All Ordinaries Index is

$$\hat{r} = \frac{\ln P_t - \ln P_0}{t} = \frac{\ln 7538.4 - \ln 557.8}{2156} = 0.001208 = \hat{R}$$

(see slide #23)

d) Fit an exponential curve to *P* with OLS.

$$P_t = P_0 e^{rt} \longrightarrow \ln P_t = \ln P_0 + rt$$

m_exp = lm(log(P) ~ t)
summary(m_exp)

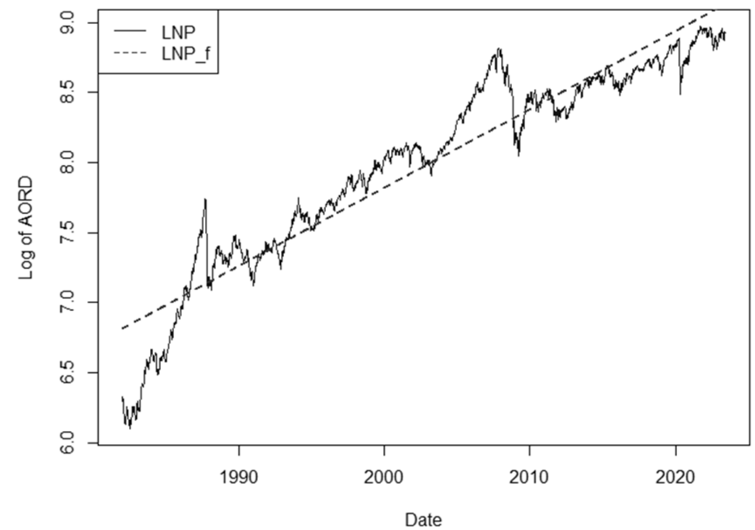
```
Call:
lm(formula = log(P) ~ t)

Residuals:
     Min       1Q   Median       3Q      Max
-0.74025 -0.12806  0.00874  0.14278  0.61113

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 6.813e+00  9.541e-03   714.1   <2e-16 ***
t           1.075e-03  7.662e-06   140.3   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2214 on 2154 degrees of freedom
Multiple R-squared:  0.9014,    Adjusted R-squared:  0.9013
F-statistic: 1.969e+04 on 1 and 2154 DF,  p-value: < 2.2e-16
```

LNP_f = ts(m_exp$fitted.values, …)
plot.ts(LNP, … col = "blue")
lines(LNP_f, type="l", lty = 2, lwd = 2,
col = "red")



At the first glance, this regression looks great as it is significant at any (reasonable) level and explains about 90% of the total sample variations of the logarithm of *P* (i.e., *LNP*).

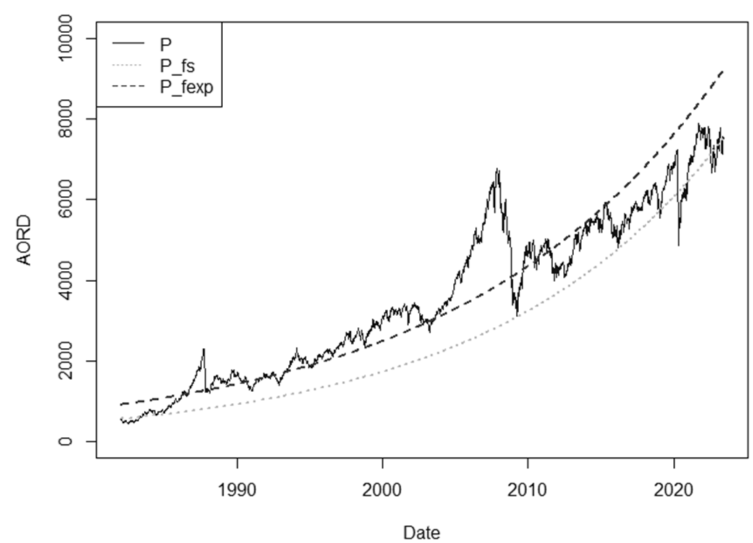
The estimate of the constant weekly log return (and that of the simple return) is *r*-hat = 0.001075 (= *R*-hat), i.e., 0.1075%, and the corresponding estimate of the annual log return is 52.14×0.1075% = 5.6051%.

In order to get estimates of *P*, we take the antilog of *LNP*,

$$\hat{P}_t = e^{\widehat{\ln P}}$$

```
P_fexp = ts(exp(LNP_f),
            frequency = 52,
            start = c(1982, 1))
```

```
plot.ts(P, xlab = "Date", ylab = "AORD",
        ylim = c(0, 10000), col = "blue")
lines(P_fc, type="l", lty = 3, lwd = 2, col = "green")
lines(P_fexp, type="l", lty = 2, lwd = 2, col = "red")
legend("topleft", legend = c("P", "P_fc", "P_fexp"),
        col = c("blue", "green", "red"), lty = c(1,3,2))
```



WHAT SHOULD YOU KNOW?

- Concepts of classical time series decomposition
- Trend, seasonal, cyclical and random components
- Price-weighted versus value-weighted stock market indices
- Simple returns and log-returns

BOARD OF FAME

Leonhard Euler (1707-1783):

Swiss mathematician, physicist, astronomer, geographer, logician and engineer

Imperial Russian Academy of Sciences

Berlin Academy

Mathematical analysis, integral calculus, trigonometric and logarithmic functions, number theory