Financial Econometrics
Financial Econometrics
Course Introduction and Overview

Contents

1 Introduction and Outline of the Course 2
2 Course Objectives 3
3 Course Structure 4
4 The Course Author 8
5 Study Materials 9
6 Assessment 9
References 18
1 Introduction and Outline of the Course

Welcome to the course on Financial Econometrics. The first objective of this course is to introduce the main econometric methods and techniques used in the analysis of issues related to finance. A course with the title *Financial Econometrics* assumes that such a field exists. However, as this quote reveals, this is far from true:

What is … financial econometrics? This simple question does not have a simple answer. The boundary of such an interdisciplinary area is always moot and any attempt to give a formal definition is unlikely to be successful. Broadly speaking, financial econometrics [aims] to study quantitative problems arising from finance. It uses statistical techniques and economic theory to address a variety of problems from finance. These include building financial models, estimation and inferences of financial models, volatility estimation, risk management, testing financial economics theory, capital asset pricing, derivative pricing, portfolio allocation, risk-adjusted returns, simulating financial systems, hedging strategies, among others (Fan, 2004: 1).

In this course, we define financial econometrics as ‘the application of statistical techniques to problems in finance’. Although econometrics is often associated with analysing economics problems such as economic growth, consumption and investment, the applications in the areas of finance have grown rapidly in the last few decades. Your textbook by Chris Brooks, *Introductory Econometrics for Finance*, lists the following examples:

1. Testing whether financial markets are weak-form informationally efficient.
2. Testing whether the CAPM or APT represent superior models for the determination of returns on risky assets.
3. Measuring and forecasting the volatility of bond returns.
4. Explaining the determinants of bond credit ratings used by the ratings agencies.
5. Modelling long-term relationships between prices and exchange rates.
6. Determining the optimal hedge ratio for a spot position in oil.
7. Testing technical trading rules to determine which makes the most money.
8. Testing the hypothesis that earnings or dividend announcements have no effect on stock prices.
9. Testing whether spot or futures markets react more rapidly to news.
10. Forecasting the correlation between the returns to the stock indices of two countries.

The above list does not include all the possibilities, and you might think of many other topics that could be added to the list.
If financial econometrics is simply the application of econometrics to finance issues, does this mean that econometric tools you have studied in previous courses are the same as those used in this course? A simple answer to this question is yes. Many of the concepts that you have encountered in the previous courses such as regression and hypothesis testing are highly relevant for this course. In fact, all the topics introduced in this course will require that you have a deep understanding of these concepts. However, the emphasis and the set of problems dealt with in finance issues are different from the economic problems you have encountered in previous courses. To start with, the nature of the data in finance issues is very different. Financial data are observed at a much higher frequency (in some instances minute-by-minute frequency). For macroeconomic data, we consider ourselves lucky if we are able to observe data on a monthly basis. Furthermore, recorded financial data such as stock market prices are those at which the transaction took place. There is no possibility for measurement error. This is in contrast to macroeconomic data, which are revised regularly.

Also the properties of financial series differ. For instance, in the course *Econometric Analysis and Applications* you spent a lot of time analysing whether the series has a unit root and devising methods to estimate models when the variables are integrated of order one. In financial econometrics, these issues are not a major concern. Although we observe prices most of the time, financial econometrics mainly deals with asset or portfolio returns. Since returns are stationary, most of the methods used in this course also apply to stationary series.

This may imply that models of financial returns are much easier to deal with. However, this is not the case. The analysis of financial data brings its own challenges. As you will see from Unit 1, financial returns possess some common properties that need to be incorporated in econometric models. For instance, returns of assets such as stocks and bonds exhibit time-varying volatility. This requires introducing new models and estimation techniques to model time varying volatility. Not only that, financial returns can exhibit asymmetry in volatility, which requires further modification of existing models. Furthermore, financial data are not normally distributed. As you have seen in previous courses, the assumption of normality has been central for estimation and hypothesis testing. Unfortunately, even in finance applications, existing econometric techniques still find it difficult to deal with models that assume non-normal distribution.

## 2 Course Objectives

After studying this course you will be able to

- define and compute measures of financial returns
- interpret sample moments of financial returns
- discuss the stylised statistical properties of asset returns
• formulate models using matrix notation
• derive the OLS estimators using matrix algebra
• use matrix algebra to analyse sources of variation of risk
• explain the principles of maximum likelihood estimation
• derive the maximum likelihood estimators and discuss their properties
• use maximum likelihood estimation, and apply the hypothesis tests available under maximum likelihood estimation
• analyse and estimate models of autoregressive, moving average, and autoregressive-moving average models
• forecast using AR, MA, and ARMA models
• apply the Box-Jenkins approach to time series models
• model and forecast volatility using autoregressive conditional heteroscedastic (ARCH) models
• estimate, interpret, and forecast with generalised autoregressive conditional heteroscedastic (GARCH) models
• extend GARCH models to analyse the asymmetric effect of shocks on volatility
• construct, estimate and interpret multivariate GARCH models
• test for spill-over of volatility between assets
• use vector autoregressive (VAR) models to analyse and interpret interaction between financial variables
• examine the impact of shocks on financial variables using impulse response analysis
• undertake tests of hypotheses and Granger causality in a VAR framework
• formulate limited dependent variable models, including logit and probit models
• estimate and interpret logit and probit models
• discuss models with multinomial linear dependent variables.

3 Course Structure

Unit 1 Statistical Properties of Financial Returns
1.1 Introduction
1.2 Calculation of Asset Returns
1.3 Stylised Facts about Financial Returns
1.4 Distribution of Asset Returns
1.5 Time Dependency
1.6 Linear Dependency across Asset Returns
Exercises | Answers to Exercises

Unit 2 Matrix Algebra, Regression and Applications in Finance
2.1 Introduction
2.2 Matrix Algebra: Some Basic Concepts and Applications
2.3 OLS Regression Using Matrix Algebra
2.4 Applications to Finance
Exercises | Answers to Exercises

Unit 3 Maximum Likelihood Estimation
3.1 Introduction
3.2 The Maximum Likelihood Function: Some Basic Ideas and Examples
3.3 The Maximum Likelihood Method: Mathematical Derivation
3.4 The Information Matrix
3.5 Usefulness and Limitations of the Maximum Likelihood Estimator
3.6 Hypothesis Testing
Exercises | Answers to Exercises

Unit 4 Univariate Time Series and Applications to Finance
4.1 Introduction
4.2 The Lag Operator
4.3 Some Key Concepts
4.4 Wold's Decomposition Theory (Optional section)
4.5 Properties of AR Processes
4.6 Properties of Moving Average Processes
4.7 Autoregressive Moving Average (ARMA) Processes
4.8 The Box-Jenkins Approach
4.9 Example: A Model of Stock Returns
4.10 Conclusions
Exercises | Answers to Exercises

Unit 5 Modelling Volatility – Conditional Heteroscedastic Models
5.1 Introduction
5.2 ARCH Models
5.3 GARCH Models
5.4 Estimation of GARCH Models
5.5 Forecasting with GARCH Model
5.6 Asymmetric GARCH Models
5.7 The GARCH-in-Mean Model
5.8 Conclusions
Exercises

Unit 6 Modelling Volatility and Correlations – Multivariate GARCH Models
6.1 Introduction
6.2 Multivariate GARCH Models
6.3 The VECH Model
6.4 The Diagonal VECH Model
6.5 The BEKK Model
6.6 The Constant Correlation Model
6.7 The Dynamic Correlation Model
6.8 Estimation of a Multivariate Model
Unit 7  Vector Autoregressive Models

7.1  Introduction
7.2  Vector Autoregressive Models
7.3  Issues in VAR
7.4  Hypothesis Testing in VAR
7.5  Example: Money Supply, Inflation and Interest Rate

Exercises | Answers to Exercises

Unit 8  Limited Dependent Variable Models

8.1  Introduction
8.2  The Linear Probability Model
8.3  The Logit Model
8.4  The Probit Model
8.5  Estimation using Maximum Likelihood
8.6  Goodness of Fit Measures
8.7  Example: Dividends, Growth and Profits
8.8  Multinomial Linear Dependent Variables
8.9  Ordered Response Linear Dependent Variable Models (optional section)

Exercises | Answers to Exercises

The objective of the course is to extend your knowledge and equip you with methods and techniques that allow you to analyse finance-related issues. This course starts by illustrating how to measure financial returns, the main variable that we try to model in financial applications. There are various definitions of returns, and Unit 1 illustrates how to compute the various types of returns. After defining financial returns, Unit 1 presents some stylised facts about the properties of financial returns. These include volatility clustering, asymmetric volatility and non-normality. Unit 1 then introduces various measures of moments of the distribution of financial returns, and how these can be computed for samples of financial returns. The material covered in this unit sets the scene for the rest of the course and thus it is important that you make yourself familiar with these concepts.

Unit 2 provides a brief introduction to the main principles of matrix algebra. In your previous courses Econometric Principles and Data Analysis and Econometric Analysis and Applications, you developed the basic regression concepts and statistical tools without referring to matrix algebra. That was essential for you to grasp some basic concepts involved in regression analysis. However, in most theoretical and practical applications, the researcher often deals with multivariate relations. As you will discover in this unit, the simplest way to tackle these multivariate relations is to switch to matrix notation. Matrix algebra eliminates the need to use summation signs and subscripts and helps present the results in a simple way. In some of the units of this course, it will be very difficult to present the proofs and results without using matrix notation. Although matrix notation simplifies the presentation of the results, the fact remains that you may be learning a new language. Learning a new language can be exciting but it is also challenging. To help you to understand and apply matrix algebra, we use matrix
algebra in some financial applications, namely the multi-factor models and portfolio theory.

Unit 3 provides a brief review of the maximum likelihood estimation method. In all the previous econometrics courses you have studied, the least squares (LS) method was used to derive the estimates of the model’s parameters and for hypothesis testing. Least Squares is just one of many estimation techniques available for econometricians. In Units 4, 5, 6 and 8 of this course, you will encounter models such as GARCH, ARMA and binary choice models that can’t be estimated by least squares. Instead, econometricians rely on maximum likelihood estimation, which is a flexible technique, more general than OLS and, under fairly general conditions, yields consistent and efficient estimates of the parameters. However, like any estimation technique, maximum likelihood is based on a certain underlying philosophy and certain principles. In Unit 3, you will be introduced to these principles and how these can be applied to derive estimates of the parameters and test hypotheses about the estimated parameters. This is one of the most challenging units, but hopefully, by using various examples, you will be able to gain a deep understanding of how the estimation method works, and you will be able to identify its strengths and weaknesses.

Unit 4 presents univariate time series models. In these types of models, a series is modelled in terms of its own past values and some disturbance terms (also known as shocks). Univariate time series models were introduced in the course Econometric Analysis and Applications. These models are different from the structural models you have studied in other courses, in the sense that these models are atheoretical – that is, they are not based on any underlying theoretical frameworks but are data driven. These models are the first building blocks for estimating financial returns and help illustrate some of the key properties of financial returns. The aim of this unit is to introduce these models, such as the autoregressive model (AR), the moving average model (MA), and a combination of these two (ARMA models).

Unit 5 presents some of the econometric methods used for modelling and forecasting volatility of asset returns. Volatility models have attracted the attention of academics and practitioners, and are widely used in many areas of finance, including models of Value-at-Risk, option pricing, and portfolio allocation. One of the stylised facts about asset returns is that the variance of the error terms is not equal at every point in time, and hence the error terms are said to suffer from heteroscedasticity. Thus, in modelling financial returns, one should consider approaches that relax the assumption of homoscedasticity. ARCH and GARCH models do exactly that. They relax the assumption of constant variance and exploit the heteroscedasticity feature to model the variance of returns over time. As you will study in this unit, ARCH models are also flexible enough to allow us to incorporate asymmetry in the volatility of financial asset returns. Please note that in this course we refer to homoscedasticity and heteroscedasticity; in Eviews, and elsewhere, you will also see the spellings homoskedasticity and heteroskedasticity.
Unit 6 extends the GARCH model from the univariate to the multivariate setting. This proves to be an important extension as it allows researchers and financial analysts to model time-varying conditional covariance and correlation between the returns of various financial assets. This new technique opened the way for many financial applications such as dynamic asset pricing models, portfolio selection, dynamic hedging, value-at-risk, and volatility transmission between assets and markets. Multivariate GARCH models also help researchers to model some of the features of asset returns, such as correlation clustering.

Unit 7 presents the Vector Autoregressive (VAR) models, which can be thought of as generalisations to the univariate time series models. VAR models represent an improvement over univariate time series models because they allow variables to be modelled not only in terms of their own lags and their own shocks, but also in terms of the lags of other variables. This provides greater flexibility and allows us to examine the dynamic interactions between a set of variables. VAR models have become very popular in the econometrics literature and are widely used in the areas of macroeconomics and finance. The tools which have developed around VAR, such as impulse response analysis, Granger causality and variance decompositions (all discussed in this unit) have become central to understanding the interaction among variables. VAR models have also been used extensively for forecasting purposes, where these models have exhibited a better performance than structural models, especially in out-of-sample forecasting.

Unit 8 deals with models in which the dependent variable i.e. the variable that needs to be explained by a set of determinants, is in fact a dummy variable. There are many cases where these models can be useful. For instance, financial analysts may be interested as to why some firms list on the stock market while others don’t; why some firms issue dividends while others don’t; and why some firms decide to raise external finance while others don’t. In all these examples, what we observe is whether a firm lists or not, issues dividends or not, or raises external finance or not. Thus, the relevant dependent variable is a dummy variable that takes the value of 1 if the event occurs, and zero if the event does not occur. Such models, known as limited dependent variable models, raise a set of estimation issues that are different from the ones you have encountered so far. The purpose of this unit is to introduce you to limited dependent models and discuss how these models can be applied to finance issues.

4 The Course Author

Bassam Fattouh graduated in Economics from the American University of Beirut in 1995. Following this, he obtained his Masters degree and PhD from the School of Oriental and African Studies, University of London, in 1999. He is a Reader in Finance and Management and academic director for the MSc in International Management for the Middle East and North Africa at
the Department for Financial and Management Studies, SOAS. He is also currently Senior Research Fellow and Director of the Oil and Middle East Programme at the Oxford Institute for Energy Studies at the University of Oxford. He has published in leading economic journals, including the *Journal of Development Economics, Economics Letters, Economic Inquiry, Macroeconomic Dynamics* and *Empirical Economics*. His research interests are mainly in the areas of finance and growth, capital structure and applied non-linear econometric modelling, as well as oil pricing systems.

## 5 Study Materials

This course mainly uses one textbook:


This textbook has been chosen for a number of reasons. It is extremely clear, contains a large number of examples and covers a lot of ground. Furthermore, it is a useful textbook to refresh your memory of some basic concepts you have studied in previous courses (especially Chapters 2, 3 and 4 of the textbook). Equally important, the textbook uses the *Eviews* package, which is quite powerful while at the same time fairly easy to use. Finally, the textbook has a very useful companion website with rich resources for students including interactive multiple choice questions, solutions to end of chapter questions, *Eviews* data and workfiles, and links to useful websites. The link for this companion website can be found at:

[http://www.cambridge.org/features/economics/brooks/student.html](http://www.cambridge.org/features/economics/brooks/student.html)

Although the textbook covers a lot of subject areas, in some units you may need to rely more heavily on the course notes and some suggested readings.

The units in the course will closely follow the presentation in the textbook. However, for some of the units, this is not feasible either because the chapter does not cover the topic at all, or covers it in a superficial way. In such cases, you may find that the course notes are more demanding than the material presented in the textbook, because the course notes analyse the issues using mathematics (though at a very basic level). This is necessary to gain a deeper understanding of the issues being considered.

Throughout this course, it is essential that you do all the readings and solve all the exercises. In this course, each idea builds on the previous ones in a logical fashion, and it is important that each idea is clear to you before you move on. You should therefore take special care not to fall behind with your schedule of studies.

## 6 Assessment

Your performance on each course is assessed through two written assignments and one examination. The assignments are written after
week four and eight of the course session and the examination is written at a local examination centre in October.

The assignment questions contain fairly detailed guidance about what is required. All assignment answers are limited to 2,500 words and are marked using marking guidelines. When you receive your grade it is accompanied by comments on your paper, including advice about how you might improve, and any clarifications about matters you may not have understood. These comments are designed to help you master the subject and to improve your skills as you progress through your programme.

The written examinations are ‘unseen’ (you will only see the paper in the exam centre) and written by hand, over a three hour period. We advise that you practice writing exams in these conditions as part of you examination preparation, as it is not something you would normally do.

You are not allowed to take in books or notes to the exam room. This means that you need to revise thoroughly in preparation for each exam. This is especially important if you have completed the course in the early part of the year, or in a previous year.

Preparing for assignments and exams

There is good advice on preparing for assignments and exams and writing them in Sections 8.2 and 8.3 of *Studying at a Distance* by Talbot. We recommend that you follow this advice.

The examinations you will sit are designed to evaluate your knowledge and skills in the subjects you have studied: they are not designed to trick you. If you have studied the course thoroughly, you will pass the exam.

Understanding assessment questions

Examination and assignment questions are set to test different knowledge and skills. Sometimes a question will contain more than one part, each part testing a different aspect of your skills and knowledge. You need to spot the key words to know what is being asked of you. Here we categorise the types of things that are asked for in assignments and exams, and the words used. The examples are from CeFiMS exam papers and assignment questions.

Definitions

Some questions mainly require you to show that you have learned some concepts, by setting out their precise meaning. Such questions are likely to be preliminary and be supplemented by more analytical questions. Generally ‘Pass marks’ are awarded if the answer only contains definitions.

They will contain such words as:

- Describe
- Define
- Examine
- Distinguish between
Course Introduction and Overview

- Compare
- Contrast
- Write notes on
- Outline
- What is meant by
- List

**Reasoning**

Other questions are designed to test your reasoning, by explaining cause and effect. Convincing explanations generally carry additional marks to basic definitions. These will include words such as:

- Interpret
- Explain
- What conditions influence
- What are the consequences of
- What are the implications of

**Judgment**

Others ask you to make a judgment, perhaps of a policy or of a course of action. They will include words like:

- Evaluate
- Critically examine
- Assess
- Do you agree that
- To what extent does

**Calculation**

Sometimes, you are asked to make a calculation, using a specified technique, where the question begins:

- Use indifference curve analysis to
- Using any economic model you know
- Calculate the standard deviation
- Test whether

It is most likely that questions that ask you to make a calculation will also ask for an application of the result, or an interpretation.

**Critique**

In many cases the question will include the word ‘critically’. This means that you are expected to look at the question from at least two points of view, offering a critique of each view and your judgment. You are expected to be critical of what you have read.

The questions may begin
Critically analyse
Critically consider
Critically assess
Critically discuss the argument that

Examine by argument
Questions that begin with ‘discuss’ are similar – they ask you to examine by argument, to debate and give reasons for and against a variety of options, for example

Discuss the advantages and disadvantages of
Discuss this statement
Discuss the view that
Discuss the arguments and debates concerning

The grading scheme
Details of the general definitions of what is expected in order to obtain a particular grade are shown below. Remember: examiners will take account of the fact that examination conditions are less conducive to polished work than the conditions in which you write your assignments. These criteria are used in grading all assignments and examinations. Note that as the criteria of each grade rises, it accumulates the elements of the grade below. Assignments awarded better marks will therefore have become comprehensive in both their depth of core skills and advanced skills.

70% and above: Distinction as for the (60–69%) below plus:

• shows clear evidence of wide and relevant reading and an engagement with the conceptual issues
• develops a sophisticated and intelligent argument
• shows a rigorous use and a sophisticated understanding of relevant source materials, balancing appropriately between factual detail and key theoretical issues. Materials are evaluated directly and their assumptions and arguments challenged and/or appraised
• shows original thinking and a willingness to take risks.

60-69%: Merit as for the (50–59%) below plus:

• shows strong evidence of critical insight and critical thinking
• shows a detailed understanding of the major factual and/or theoretical issues and directly engages with the relevant literature on the topic
• develops a focussed and clear argument and articulates clearly and convincingly a sustained train of logical thought
• shows clear evidence of planning and appropriate choice of sources and methodology.

50–59%: Pass below Merit (50% = pass mark)

• shows a reasonable understanding of the major factual and/or theoretical issues involved
• shows evidence of planning and selection from appropriate sources,
• demonstrates some knowledge of the literature
• the text shows, in places, examples of a clear train of thought or argument
• the text is introduced and concludes appropriately.

45–49%: Marginal Failure

• shows some awareness and understanding of the factual or theoretical issues, but with little development
• misunderstandings are evident
• shows some evidence of planning, although irrelevant/unrelated material or arguments are included.

0–44%: Clear Failure

• fails to answer the question or to develop an argument that relates to the question set
• does not engage with the relevant literature or demonstrate a knowledge of the key issues
• contains clear conceptual or factual errors or misunderstandings.

Specimen exam paper

Your final examination will be very similar to the Specimen Exam Paper that follows. It will have the same structure and style and the range of question will be comparable.

CeFiMS does not provide past papers or model answers to papers. Our courses are continuously updated and past papers will not be a reliable guide to current and future examinations. The specimen exam paper is designed to be relevant to reflect the exam that will be set on the current edition of the course.

Further information

The OSC will have documentation and information on each year’s examination registration and administration process. If you still have questions, both academics and administrators are available to answer queries.

The Regulations are also available at www.cefims.ac.uk/regulations.shtml, setting out the rules by which exams are governed.
This is a specimen examination paper designed to show you the type of examination you will have at the end of the year for Financial Econometrics. The number of questions and the structure of the examination will be the same but the wording and the requirements of each question will be different. Best wishes for success in your final examination.

The examination must be completed in THREE hours.

Answer THREE questions. The examiners give equal weight to each question; therefore, you are advised to distribute your time approximately equally between three questions.

Do not remove this paper from the examination room.

It must be attached to your answer book at the end of the examination.
1. Answer all parts of the question.
Consider the following GARCH(1,1) model
\[ r_t = \mu + u_t, \quad u_t \sim N(0,1) \]
\[ \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \]
where \( r_t \) is a daily stock return series.

a) Explain carefully how you would estimate this model.
b) Interpret the estimated coefficients of the model.
c) Explain the difference between the conditional variance and the unconditional variance. Calculate the unconditional variance for the model. Under what conditions will unconditional variance be stationary?
d) Describe two extensions to the original GARCH model. What additional characteristics of financial data might they be able to capture?

2. Answer all parts of the question.
Suppose that you are interested in modelling the correlation between the returns of the British Airways stock and the returns on crude oil.
a) Write down a constant correlation problem for this model.
b) Explain carefully how you would estimate the above model.
c) Discuss the values of the coefficients estimates that you would expect.
d) Discuss two alternative approaches to correlation modelling. What are the strengths and weaknesses of multivariate GARCH model compared to the alternatives you propose?

3. Answer all parts of the question.
a) Explain why the linear model is inadequate as a specification for the limited dependent variable estimation.
b) How does the logit model overcome the problem in part (a)?
c) Describe the intuition behind the maximum likelihood estimation technique used for limited dependent variable model.
d) How can we measure whether a logit model that has been estimated fits the data well or not?
4 Answer all parts of the question.
   a) Explain the underlying philosophy behind maximum likelihood estimation.
   b) OLS and maximum likelihood are used to estimate the parameters of a standard linear regression model. Will they give the same estimate? Explain your answer.
   c) Discuss the main advantages and limitations of the maximum likelihood method.
   d) Describe one hypothesis testing procedure that is available under the maximum likelihood estimation.

5 Answer all parts of the question.
   a) Discuss three stylised features of financial data.
   b) Can any of these features be modelled using linear time series models? Explain your answer.
   c) Explain the moments of the random variable. How can you estimate these in empirical applications?
   d) Explain carefully the Bera-Jarque test, stating clearly the null and alternative hypothesis. In case of financial data, do you expect the Bera-Jarque test to accept or reject the null? Explain your answer.

6 Answer all parts of the question.
Consider the following three models, which a researcher suggests might be a reasonable model of stock market prices.

\[ y_t = y_{t-1} + u_t \]  
\[ y_t = 0.5 y_{t-1} + u_t \]  
\[ y_t = 0.8 u_{t-1} + u_t \]

   a) What classes of models are these examples of?
   b) What would the autocorrelation function for each of these processes look like? (Don’t calculate the acf; simply consider the shape.)
   c) Describe the steps that Box and Jenkins suggested should be involved in constructing the above models?
   d) What procedure may be used to estimate the parameter in equation 3? Explain how such a procedure works and why OLS is no appropriate.
7. Answer all parts of the question.
   a) Explain why VAR models have become popular for application in economics and finance, relative to structural models derived from some underlying theory.
   b) Define carefully the following terms
      i. structural VAR model
      ii. standard VAR model.
   c) Describe and evaluate two methods for determining the appropriate lag lengths in VARs.
   d) Explain carefully the Granger Causality test in a multivariate framework.

8. Answer all parts of the question.
   Consider the following model where the dependent variable (say stock returns) is assumed to be dependent on $k$ explanatory variables such that
   \[ y_1 = \beta_1 + \beta_2 x_{2,1} + \beta_3 x_{3,1} + \beta_4 x_{4,1} + \ldots + \beta_k x_{k,1} + u_1 \]
   \[ y_2 = \beta_1 + \beta_2 x_{2,2} + \beta_3 x_{3,2} + \beta_4 x_{4,2} + \ldots + \beta_k x_{k,2} + u_2 \]
   \[ y_T = \beta_1 + \beta_2 x_{2,T} + \beta_3 x_{3,T} + \beta_4 x_{4,T} + \ldots + \beta_k x_{k,T} + u_T \]
   a) Write the above model in matrix form.
   b) Estimate the parameters of the model in part (a) using the method of least squares.
   c) What is mean by multifactor models? How are these models used in finance?
   d) Consider a two factor model such that
      \[ y = \alpha + X\alpha + u \]
      where $y$ is the vector of returns, $\alpha$ is the vector of ones and $X$ is the matrix of the factors. Calculate the expected value and the variance of returns. Comment on the results.

[END OF EXAMINATION]
References


Financial Econometrics

Unit 1 Statistical Properties of Financial Returns

Contents

1.1 Introduction 3
1.2 Calculation of Asset Returns 4
1.3 Stylised Facts about Financial Returns 10
1.4 Distribution of Asset Returns 11
1.5 Time Dependency 15
1.6 Linear Dependency across Asset Returns 17
1.7 Conclusion 20

References 20
Exercises 21
Answers to Exercises 22
Unit Content

Unit 1 explains how to calculate returns on financial assets, and considers various stylised facts (common statistical properties) concerning financial returns. The unit then analyses the distribution of returns, and, using examples, tests whether the various returns follow the Normal distribution. Following that, you will study an analysis of time dependency, considering serial correlation in returns, serial correlation in volatility and asymmetry of volatility. An important finding is that time dependency can occur at more than one level (often time dependency exists in terms of the variance of the return but not the mean), and models of financial returns should take this into account.

Learning Outcomes

When you have completed your study of this unit, you will be able to

- define and compute the various measures of financial returns, including the simple return, gross return, multi-period returns, continuously compounded returns
- calculate the sample moments of financial returns, including the skewness and kurtosis of financial returns, using Eviews
- explain and discuss some of the stylised statistical properties of asset returns
- analyse and appreciate the issue of time dependency in asset returns
- analyse the linear dependence across financial assets.

Reading for Unit 1


There are optional readings from Chris Brooks (2008) Introductory Econometrics for Finance.
1.1 Introduction

The main purpose of this unit is to describe and analyse some of the properties of returns on financial assets. Although financial analysts often observe prices on their screens such as stock prices, commodity prices, bond prices and exchange rates, the main objective of financial econometrics is to analyse financial returns. The focus on returns has many advantages. Being computed as a difference between prices over a certain horizon, financial returns are stationary. This allows us to apply many of the standard calculation methods, summary statistics, and the standard econometric techniques you have studied before. Furthermore, returns can be easily compared across assets since they are scale free. For instance, you could compare the annual return of an investment in stocks with an investment in a bond. Finally, as you will see in this unit, by focusing on financial returns it is possible to describe some common statistical properties of asset returns. These common features can be useful in modelling the time series properties of financial returns.

This unit starts by illustrating how to measure financial returns, the main variable that we try to model in financial applications. There are various definitions of returns such as simple returns, gross returns, multi-period returns, log returns, and so on. It is important from the start to be clear on how to compute the various types of returns. It is worth stressing that although financial returns are scale free, they should always be defined with respect to a certain time interval. This will be illustrated using many examples.

After defining financial returns, we present some stylised facts about the properties of financial returns. As noted by Cont and Tankov (2004: 209–10),

> After all, why should properties of corn futures be similar to those of IBM shares or the Dollar/Yen exchange rate? Nevertheless, the result of more than half a century of empirical studies on financial time series indicates that this is the case if one examines their properties from a statistical point of view. The seemingly random variations of asset prices do share some quite nontrivial statistical properties. Such properties, common across a wide range of instruments, markets and time periods are called stylized empirical facts.

In this unit, we focus on some of these properties, mainly the time dependency properties, volatility clustering, asymmetric volatility, non-normality and cross-correlations across assets. But before doing so, it is important to refresh your memory about the various measures of moments of the distribution of a random variable and how these can be computed for samples of financial returns. This has an additional advantage since it will allow you to learn how to derive these measures using simple commands in Eviews.

The reading of this unit will be based on Chapter 2 of the textbook by Eric Jondeau, Ser-Huang Poon and Michael Rockinger, Financial Modelling under Non-Gaussian Distributions, which is reprinted in your Reader.
Although this reading is extracted from an advanced econometrics textbook, it sets out the issues in a clear and an insightful way. The outline of this unit follows very closely that of the reading. However, the course notes will discuss some of the issues in more detail, and will try to reproduce some of the results using new datasets and using Eviews. It is important to note that many of the issues introduced in this unit will be revisited in other units and thus one of the purposes of this unit is to set the scene for the rest of the course.

### 1.2 Calculation of Asset Returns

Although in financial markets we mostly observe asset prices such as share prices or commodity prices, in empirical applications we often work with returns. One major reason for dealing with returns is that while prices are non-stationary (i.e. asset prices contain a unit root), asset returns are stationary. Since the course deals heavily with analysing and estimating asset return equations, it is worth spending some time defining returns and highlighting some of stylised facts about financial returns.

#### 1.2.1 Simple returns

There are various definitions of returns. One such definition is the simple return. Let $P_t$ be the price of an asset at time $t$ and let $P_{t-1}$ be the price of the asset at time $t - 1$. Assuming that the financial asset does not pay any dividends, then the one-period (for instance, one-day, one-week, one-month or one-year) simple net return denoted as $R_t$ is given by the following equation

$$ R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{1.1} $$

Writing

$$ \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{P_t}{P_{t-1}} - 1 $$

one can define the one-period simple gross return as

$$ 1 + R_t = \frac{P_t}{P_{t-1}} \tag{1.2} $$

The left-hand side of the equation is also known as the discrete compounding factor. This is the case since we could write Equation (1.2) as

$$ P_t = (1 + R_t) P_{t-1} \tag{1.3} $$

It is important to stress that returns should always be defined with respect to a certain time interval. For instance, a statement such as ‘the investment achieved a return of 20%’ is meaningless unless we specify the horizon in which this return has been achieved. Thus, the above sentence should be qualified to include the time horizon, such as ‘the investment achieved
a monthly return of 20%’ or ‘the investment achieved an annual return of 20%’.

Review Question
Consider a one-month investment in a BMW share. You bought the stock in period \( t-1 \) at $90 and sold it in period \( t \) for $100. Calculate the simple net return and the gross return of holding the investment over this one-month period.

The one-month simple net return is
\[
R_t = \frac{100 - 90}{90} = 11.11\%
\]
The one-month simple gross return is given by
\[
1 + R_t = \frac{100}{90} = 111.11\%
\]

1.2.2 Multiperiod returns
Suppose that you hold a financial asset from period \( t-k \) to \( t \), then the multiperiod simple net return denoted as \( R_t(k) \) is given by the following
\[
R_t(k) = \frac{P_t - P_{t-k}}{P_{t-k}} \tag{1.4}
\]

For instance, assume that you hold the financial asset for two periods from \( t-2 \) to \( t \) then the two-period net simple return is given by
\[
R_t(2) = \frac{P_t - P_{t-2}}{P_{t-2}} = \frac{P_t}{P_{t-2}} - 1
\]
Writing
\[
\frac{P_t}{P_{t-2}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}}
\]
the two-period simple net return can be written as
\[
R_t(2) = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} - 1
\]
which yields
\[
R_t(2) = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} - 1 = (1 + R_t) \times (1 + R_{t-1}) - 1
\]
or
\[
1 + R_t(2) = (1 + R_t) \times (1 + R_{t-1})
\]
Notice that the simple two-period gross return is a geometric sum of the two one-period simple gross returns. Thus, adding two simple one-period gross returns does not yield the two-period return.
More generally, the $k$-period gross return can be written as

$$1 + R_t(k) = \left(1 + R_t\right) \times \ldots \times \left(1 + R_{t-k}\right)$$  \hfill (1.5)

**Review Question**

Continue with the above example, but suppose now that you hold the asset for two months and in month $t-2$ the price was $50. Calculate the two-month net return and gross return.

The two-month net return is given by

$$R_t(2) = \frac{P_t - P_{t-2}}{P_{t-2}} = \frac{100 - 50}{50} = 100\%$$

The two-month gross return is given by

$$1 + R_t(2) = \left(1 + R_t\right) \times \left(1 + R_{t-1}\right)$$

where

$$1 + R_t = \frac{100}{90} = 1.11$$

$$1 + R_{t-1} = \frac{1 + \ln \frac{P_t}{P_{t-1}}}{\ln \frac{P_t}{P_{t-1}}} = \frac{90}{50} = 1.80$$

Substituting the values in the above equation (without rounding) yields

$$1 + R_t(2) = 1.11 \times 1.80 = 200\%$$

### 1.2.3 Portfolio return

The simple net return for an $N$-asset portfolio, denoted as $R_{p,t}$, is just the weighted average of individual simple returns. Thus,

$$R_{p,t} = \sum_{i=1}^{N} w_i R_{i,t}$$ \hfill (1.6)

where $w_i$ is the weight of asset $i$ in the portfolio and $N$ is the number of assets in the portfolio. This is an extremely useful property for simple returns, and thus when dealing with portfolio analysis, it is easier to calculate simple returns.

### 1.2.4 Log returns

In this course, we will base most of our examples on continuously compounded returns. The continuously compounded one-period return (or log return) denoted as $r_t$ is given by

$$r_t = \ln \left(1 + R_t\right) = \ln \left(\frac{P_t}{P_{t-1}}\right) = \ln P_t - \ln P_{t-1}$$ \hfill (1.7)
where $\ln$ is the natural log function. Another way to express the above function is as follows

$$\exp(r_t) = \frac{P_t}{P_{t-1}}$$  \hspace{1cm} (1.8)

The left-hand side of equation (1.8) refers to the continuously compounding factor since equation (1.8) can be written as

$$P_t = \exp(r_t)P_{t-1}$$  \hspace{1cm} (1.9)

### 1.2.5 Multiperiod log returns

The main advantage of using log returns is that the multiperiod return is simply the sum of one-period returns. In other words,

$$r_t(k) = \sum_{j=0}^{k-1} r_{t-j}$$  \hspace{1cm} (1.10)

This is a very useful property, which is extremely helpful in practical applications, as you will see in the next exercise.

---

**Review Question**

Table 1.1 contains monthly share prices (adjusted for splits and dividends) for Barclays Bank from December 2007 to December 2008 and the monthly log returns. The data were obtained from yahoo finance. Using equation (1.7), check that you can calculate the one-month log-returns. Using equation (1.10), check that you can calculate the annualised continuously compounded returns for 2008.

<table>
<thead>
<tr>
<th>Date</th>
<th>Share Price</th>
<th>Monthly Log Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 2007</td>
<td>504</td>
<td></td>
</tr>
<tr>
<td>January 2008</td>
<td>470</td>
<td>−0.069843573</td>
</tr>
<tr>
<td>February 2008</td>
<td>477.25</td>
<td>0.015307768</td>
</tr>
<tr>
<td>March 2008</td>
<td>453</td>
<td>−0.052148337</td>
</tr>
<tr>
<td>April 2008</td>
<td>456.5</td>
<td>0.007696575</td>
</tr>
<tr>
<td>May 2008</td>
<td>375</td>
<td>−0.196662674</td>
</tr>
<tr>
<td>June 2008</td>
<td>291.5</td>
<td>−0.25188602</td>
</tr>
<tr>
<td>July 2008</td>
<td>338</td>
<td>0.14800589</td>
</tr>
<tr>
<td>August 2008</td>
<td>353</td>
<td>0.043422161</td>
</tr>
<tr>
<td>September 2008</td>
<td>326.5</td>
<td>−0.078038108</td>
</tr>
<tr>
<td>October 2008</td>
<td>178.9</td>
<td>−0.601602958</td>
</tr>
<tr>
<td>November 2008</td>
<td>169.4</td>
<td>−0.054564208</td>
</tr>
<tr>
<td>December 2008</td>
<td>153.4</td>
<td>−0.099213893</td>
</tr>
</tbody>
</table>

**Annualised continuously compounded return**

To calculate the annualised continuously compounded returns for 2008, you simply add the monthly log returns to obtain $−1.1895$. Alternatively, you can calculate the average monthly return ($−0.09913$) and then multiply it by 12 to obtain the annualised...
continuously compounded returns (–1.1895); in this example, this step might seem pointless (dividing by 12 observations and then multiplying by 12 months), but it is required if you do not have 12 observations.

1.2.6 **Real log returns**

So far, we have only considered nominal returns. In some practical applications, we may also be interested in real returns (i.e. nominal returns adjusted for the inflation rate). The log returns are quite useful in calculating real returns.

Calculating the real return involves two steps. In the first step, you deflate the share price by the general price level (usually the Consumer Price Index, CPI). In the second step, you calculate the return using the same methods as applied above. As an example, consider $P_t$ the price of the share at time $t$ and $CPI_t$ is the consumer price at time $t$. The real share price is given by the following

$$P_t^{\text{Real}} = \frac{P_t}{CPI_t} \quad (1.11)$$

The one-period simple real return is computed as

$$R_t^{\text{Real}} = \frac{P_t^{\text{Real}} - P_{t-1}^{\text{Real}}}{P_{t-1}^{\text{Real}}} = \left( \frac{P_t}{CPI_t} - \frac{P_{t-1}}{CPI_{t-1}} \right) + \frac{P_{t-1}}{CPI_{t-1}} - \frac{CPI_t}{CPI_{t-1}} - 1 \quad (1.12)$$

The continuously compounded one-period real return denoted as $r_t^{\text{Real}}$ is given by the following

$$r_t^{\text{Real}} = \ln \left( 1 + R_t^{\text{Real}} \right) = \ln \left( \frac{P_t}{P_{t-1}} \times \frac{CPI_t}{CPI_{t-1}} \right) \quad (1.13)$$

Using the log properties, equation (1.13) can be written as

$$r_t^{\text{Real}} = \left( \ln \left( P_t \right) - \ln \left( P_{t-1} \right) \right) - \left( \ln \left( CPI_t \right) - \ln \left( CPI_{t-1} \right) \right) \quad (1.14)$$

The first term on the right-hand side is simply the log return, while the second term is the one-period continuously compounded inflation rate ($\pi_t$) i.e. equation (1.14) can be written as

$$r_t^{\text{Real}} = r_t - \pi_t \quad (1.15)$$

**Review Question**

Table 1.2 contains monthly data on the New York Stock Exchange Price Index and the monthly Consumer Price Index (CPI) for the US. Using equation (1.14), check that you can calculate the monthly real rate of return.
Table 1.2  New York Stock Exchange Price Index and CPI, December 2007–December 2008

<table>
<thead>
<tr>
<th>Date</th>
<th>NYSE Price Index</th>
<th>$r_t$</th>
<th>CPI</th>
<th>$\pi_t$</th>
<th>$r_t^{\text{Real}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 2007</td>
<td>9740.32</td>
<td>211.737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>January 2008</td>
<td>9126.16</td>
<td>-0.065129</td>
<td>212.495</td>
<td>0.0035735</td>
<td>-0.0687025</td>
</tr>
<tr>
<td>February 2008</td>
<td>8962.46</td>
<td>-0.0181003</td>
<td>212.86</td>
<td>0.0017162</td>
<td>-0.0198165</td>
</tr>
<tr>
<td>March 2008</td>
<td>8797.29</td>
<td>-0.018601</td>
<td>213.667</td>
<td>0.0037841</td>
<td>-0.0223851</td>
</tr>
<tr>
<td>April 2008</td>
<td>9299.6</td>
<td>0.05552767</td>
<td>213.997</td>
<td>0.0015433</td>
<td>0.0539844</td>
</tr>
<tr>
<td>May 2008</td>
<td>9401.08</td>
<td>0.01085319</td>
<td>215.044</td>
<td>0.0048807</td>
<td>0.0059725</td>
</tr>
<tr>
<td>June 2008</td>
<td>8660.48</td>
<td>-0.0820544</td>
<td>217.034</td>
<td>0.0092114</td>
<td>-0.0912658</td>
</tr>
<tr>
<td>July 2008</td>
<td>8438.64</td>
<td>-0.025949</td>
<td>218.61</td>
<td>0.0072353</td>
<td>-0.0331843</td>
</tr>
<tr>
<td>August 2008</td>
<td>8382.08</td>
<td>-0.0067251</td>
<td>218.576</td>
<td>-0.0001555</td>
<td>-0.0065695</td>
</tr>
<tr>
<td>September 2008</td>
<td>7532.8</td>
<td>-0.1068293</td>
<td>218.675</td>
<td>0.0004528</td>
<td>-0.1072821</td>
</tr>
<tr>
<td>October 2008</td>
<td>6061.09</td>
<td>-0.2173772</td>
<td>216.889</td>
<td>-0.0082009</td>
<td>-0.2091763</td>
</tr>
<tr>
<td>November 2008</td>
<td>5599.3</td>
<td>-0.0792481</td>
<td>213.263</td>
<td>-0.0168596</td>
<td>-0.0623885</td>
</tr>
<tr>
<td>December 2008</td>
<td>5757.05</td>
<td>0.0277836</td>
<td>211.577</td>
<td>-0.0079371</td>
<td>0.0357207</td>
</tr>
</tbody>
</table>

Annualised continuously compounded real return: $-0.5250928$

As can be seen from Table 1.2, the monthly real rate of return is simply the monthly log return minus the one month continuously compounded inflation. To calculate the annualised real rate of return for 2008, you can simply add the real monthly log returns.

1.2.7 Log portfolio return

The main disadvantage of using log returns is that the log return of a portfolio of assets cannot be written as the weighted average of individual simple returns. In fact, the portfolio log return denoted as $r_{p,t}$ is given by

$$
    r_{p,t} = \ln \left( 1 + R_{p,t} \right) = \ln \left( 1 + \sum_{i=1}^{N} w_i R_{i,t} \right) \neq \sum_{i=1}^{N} w_i r_{i,t}
$$

(1.16)

This is the case because the log of a sum is different from the sum of logs. In your next reading, the authors claim that this problem is usually considered minor in empirical applications. This is true to some extent, especially when returns are measured over short intervals of time. In such cases,

$$
    r_{p,t} \approx \sum_{i=1}^{N} w_i r_{i,t}
$$

(1.17)

However, it is not advisable to use this approximation, and when you need to construct portfolio returns, it is better to use simple returns. In this course we will be mainly examining the behaviour of asset returns over time, and not portfolio returns, so we will rely heavily on log returns.
Reading

I would like you now to read Section 2.1 of the chapter by Jondeau, Poon and Rockinger. If you are unsure about how to calculate any of the above returns, perhaps at this stage it would also be useful to revise the properties of logarithms. Appendix A3 of your textbook by Chris Brooks (pp.608–09) provides a quick review of the properties of logarithms. You could also read pp 7–9 of Brooks, which covers simple returns, log returns, and log returns of a portfolio.

1.3 Stylised Facts about Financial Returns

Although different assets such as stocks, bonds or commodities behave differently and are unlikely to be affected by the same set of information or events, the vast empirical literature on financial time series over the last few decades has revealed that financial asset returns possess some common statistical properties. These properties are often referred to as stylised facts. In what follows, we choose the most important stylised facts as listed by Cont and Tankov (2004: 211).

1 Absence of autocorrelations: (linear) autocorrelations of asset returns are often insignificant, except for very small intraday time scales (≈ 20 minutes) for which microstructure effects come into play.

2 Heavy tails: the (unconditional) distribution of returns seems to display a power-law or Pareto-like tail, with a tail index which is finite, higher than two and less than five for most data sets studied. In particular, this excludes stable laws with infinite variance and the normal distribution. However the precise form of the tails is difficult to determine.

3 Gain/loss asymmetry: one observes large drawdowns in stock prices and stock index values but not equally large upward movements.

4 Aggregational Gaussianity: as one increases the time scale $\Delta t$ over which returns are calculated, their distribution looks more and more like a normal distribution. In particular, the shape of the distribution is not the same at different time scales.

6 Volatility clustering: different measures of volatility display a positive autocorrelation over several days, which quantifies the fact that high-volatility events tend to cluster in time.

7 Conditional heavy tails: even after correcting returns for volatility clustering (e.g. via GARCH-type models), the residual time series still exhibit heavy tails. However, the tails are less heavy than in the unconditional distribution of returns.

9 Leverage effect: most measures of volatility of an asset are negatively correlated with the returns of that asset.

10 Volume/volatility correlation: trading volume is correlated with all measures of volatility.
1.4 Distribution of Asset Returns

In the rest of the course, we will analyse some of these properties in detail, and discuss how different models try to incorporate these features. But even at this early stage, it is worth illustrating some of these stylised facts using data on stock market indices. However, before doing so, it would be useful to refresh your memory about the moments of a random variable, and then show you how these can be used to illustrate the properties of financial returns.

1.4.1 Moments of a random variable

Denote by the random variable \( X \) the log return of a financial asset. As you have seen in previous courses, the cumulative distribution function for the random variable can be defined as

\[
F(X) = \Pr[X \leq x] = \int_{-\infty}^{x} f_X(u)du
\]  

(1.18)

where \( f_X \) is the probability density function (pdf). The un-centred moments of the random variable \( X \) are defined as

\[
m_k = E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x)dx \quad \text{for } k = 1, 2, \ldots
\]  

(1.19)

Although the above equations seem complex, their interpretation is quite straightforward. When \( k = 1 \), you obtain the first un-centred moment of the random variable, which is simply the mean of the random variable i.e.

\[
m_1 = E[X] = \mu
\]  

(1.20)

The centred first moment equals zero. When \( k = 2 \), we can obtain the second centred moment of the random variable, which is simply the variance i.e.

\[
m_2 = E[(X-\mu)^2] = \nu(X) = \sigma^2
\]  

(1.21)

When \( k = 3 \), we obtain the skewness of the random variable, and when \( k = 4 \), we obtain the kurtosis of the series. The (standardised) skewness of the series, denoted as \( s \), is defined as

\[
s = Sk[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]
\]  

(1.22)

The (standardised) kurtosis of the series, denoted as \( \kappa \), is defined as

\[
\kappa = Ku[X] = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]
\]  

(1.23)

The skewness and kurtosis of the series are important for understanding financial returns. Skewness measures the asymmetry of the distribution of financial returns. When it is positive, it indicates that large positive realisations of \( X \) are more likely. When it is negative, it indicates that large negative
realisations of $X$ are more likely. Kurtosis, on the other hand, measures the thickness of the tails of the distribution. In particular, it measures the tail thickness in relation to the normal distribution (for a normal distribution kurtosis equals 3, so excess kurtosis is measured by $\kappa - 3$). Remember from the above discussion that one of the stylised facts is that financial returns have heavy tails, and that these heavy tails persist even after correcting for volatility clustering.

### 1.4.2 Empirical moments

In practical application, we need to consider empirical measures for the above moments. In previous courses, you have studied how to compute these moments, but it is worth reviewing these very quickly. Consider a time series of realised asset returns $r_t$, $t = 1, \ldots, T$. The widely used measure of location is the sample mean, which is given by the following equation:

$$\bar{r} = \hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} r_t$$  \hspace{1cm} (1.24)

Variance is the most widely used measure for dispersion, and is given by the following equation

$$\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^{T} (r_t - \bar{r})^2$$  \hspace{1cm} (1.25)

In financial applications, the square root of the variance is often used to measure volatility. Another useful measure of dispersion is the Mean Absolute Deviation (MAD), which is given by

$$MAD = \frac{1}{T} \sum_{t=1}^{T} |r_t - \bar{r}|$$  \hspace{1cm} (1.26)

The sample skewness can be computed using the following equation

$$\hat{s} = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{r_t - \bar{r}}{\sigma}\right)^3$$  \hspace{1cm} (1.27)

The sample kurtosis can be computed using the following equation

$$\hat{k} = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{r_t - \bar{r}}{\sigma}\right)^4$$  \hspace{1cm} (1.28)

The above measures are known as summary statistics. Under the assumption that financial returns are normal, we have the following asymptotic results:

$$\sqrt{T} \left(\hat{\mu} - \mu\right) \sim N\left(0, \sigma^2\right)$$  \hspace{1cm} (1.29)

$$\sqrt{T} \left(\hat{\sigma}^2 - \sigma^2\right) \sim N\left(0, 2\sigma^4\right)$$  \hspace{1cm} (1.30)

$$\sqrt{T} \ \hat{s} \sim N\left(0, 6\right)$$  \hspace{1cm} (1.31)
\[
\sqrt{T} \left( \hat{\kappa} - 3 \right) \sim N \left( 0, 24 \right) \quad (1.32)
\]

In fact, based on the results in (1.31) and (1.32), one can derive a statistic to test the hypothesis of normality, known as the Bera-Jarque test, defined as

\[
JB = T \left[ \frac{s^2}{6} + \frac{(\hat{\kappa} - 3)^2}{24} \right] \quad (1.33)
\]

The test is distributed asymptotically as \( \chi^2(2) \) under the hypothesis that the distribution is normal. A large value of the J-B statistic implies that we can reject the null hypothesis that the returns are normally distributed.\(^1\)

**Example**

Let’s now use these measures to illustrate some of the properties of financial returns. The example concerns daily, weekly, monthly and annual data for the FTSE stock price index from April 1990 to January 2009. For each of the series we calculate the corresponding one-period log returns (daily log returns for daily data, weekly log returns for weekly data, and monthly log returns for monthly data). For each of the series, we calculate the mean, standard deviation, the skewness and kurtosis. The summary statistics are shown in the Figures below.

For daily data (Figure 1.1), the mean of the daily log return is 0.016% and the standard deviation is 1.15%, which is quite high.

![Figure 1.1 Daily Log Returns, FTSE](image)

**Note:** Eviews automatically provides these summary statistics. Once you have opened the file go to:

1. Quick
2. Scroll down to ‘Series Statistics’
3. Then choose ‘Histogram and Stats’.

\(^1\) It is important to stress that the JB test applies to only large samples, as explained in your reading.
Notice the maximum and minimum, which range from 9% to –9%. As expected, the daily index return has high sample kurtosis of 9.60, a clear sign of fat tails. The daily index return is slightly negatively skewed. The J-B test strongly rejects the null hypothesis of normality.

For weekly data (Figure 1.2), the mean of the log return is 0.078% and the standard deviation is 2.4%. Notice the wide maximum and minimum range of 12% to -23%. Interestingly, the weekly index return has a very high sample kurtosis of 14.35, which is higher than the kurtosis for daily returns. The weekly returns also have higher negative skewness than daily returns. Again, the J-B test strongly rejects the null hypothesis of normality.

Figure 1.2  Weekly Log Returns, FTSE

![Weekly Log Returns, FTSE](image1)

Finally, we report data for monthly returns (Figure 1.3).

Figure 1.3  Monthly Log Returns, FTSE

![Monthly Log Returns, FTSE](image2)

The monthly log return has a relatively low kurtosis of 3.67. This is expected for monthly data. The monthly returns still exhibit negative skewness. Again, the J-B test strongly rejects the null hypothesis of normality of returns.
1.5 Time Dependency

As noted above, one of the stylised facts is that autocorrelations of asset returns are often insignificant i.e. asset returns exhibit no time dependency. However, it is important to note that time dependency can occur at several levels. In what follows, we refer to three levels of dependency: Serial correlation in returns, serial correlation in squared return, and volatility asymmetry.

1.5.1 Serial correlation in returns

Here, we are interested in testing the null hypothesis that the first $p$ returns are not serially correlated. Remember from previous courses that a measure of autocorrelation of returns of order $j$ is given by the following:

$$
\hat{\rho}_j = \frac{\sum_{t=j+1}^{T} (r_t - \bar{r})(r_{t-j} - \bar{r})}{\sum_{t=1}^{T} (r_t - \bar{r})^2}
$$

(1.34)

Unit 4 will use test statistics such as the Ljung-Box $Q$ statistic to test the significance of autocorrelations, and will suggest ways to estimate models of financial returns. However, we will use this statistic in the example that follows, and it is calculated as follows:

$$
Q_p = T(T + 2) \sum_{j=1}^{p} \frac{1}{T-j} \hat{\rho}_j^2
$$

(1.35)

It is asymptotically distributed as $\chi^2$ with $p$ degrees of freedom, under the null hypothesis of no correlation. As discussed in the stylised facts, autocorrelations of asset returns are often insignificant and hence there is little time dependency in asset returns. However, this stylised fact cannot be generalised. Depending on the time horizon being used, one could find weak evidence of serial correlation in some asset returns.

1.5.2 Serial correlation in volatility

To test for dependency in volatility, we need to construct models that generate time-varying volatility measures. ARCH, GARCH and their family of models do exactly that. In Unit 5, we will introduce these models as well as ways to test for serial dependence in volatility. To anticipate the discussion in Unit 5, we could use the Ljung-Box $Q$ statistic to test for serial correlation in squared returns and absolute returns. Most empirical evidence suggests that there is a strong evidence of serial correlations in squared returns and absolute returns, especially for daily and weekly data as shown in Table 2.4 of your reading. In other words, large returns of either sign tend to be followed...
by large returns of either sign or the volatility of returns tends to be serially correlated. This is often referred to in the literature as volatility clustering.

1.5.3 **Volatility asymmetry**

One important feature of financial returns is that volatility exhibits asymmetric behaviour. In particular, there is wide empirical evidence that volatility is more affected by negative returns than positive returns. In Unit 5, we will show you how these ARCH and GARCH models can be modified to take asymmetric volatility into account. Table 2.5 of your reading shows parameter estimates of volatility asymmetry for the various stock market indices.

---

**Reading**

I would like you now to read Section 2.3 in Jondeau, Poon and Rockinger. Don’t worry if you don’t understand all of these equations. These will become clear in Units 4 and 5. The main lessons I want you take from this section are as follows.

- Time dependency can occur at more than one level, and for financial returns time dependency often occurs at the second moment (the variance) and not the first moment (the mean);
- Therefore, it is important to construct models of time varying volatility for financial returns and devise statistics to test for the correlation at higher moments;
- Volatility of financial returns may exhibit asymmetric behaviour and these need to be accounted for in empirical models.

---

**Example**

Perhaps the best way to appreciate the issue of time dependency is to consider again the monthly log return of the FTSE. The serial correlation at order 1 to 6 and the corresponding Ljung-Box $Q$ statistic are given in Table 1.3. In Unit 4, you will learn how to derive such a table using *Eviews*, but for now it is important to understand the intuition. As can be seen from this table, there is no evidence of serial correlation in the monthly log returns. The $Q$ statistic does not reject the null of no serial correlation at the various lags.

**Table 1.3 Monthly Log Returns, FTSE**

<table>
<thead>
<tr>
<th>Lag</th>
<th>Autocorrelation</th>
<th>Ljung-Box $Q$ Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.081</td>
<td>1.5497</td>
<td>0.213</td>
</tr>
<tr>
<td>2</td>
<td>−0.071</td>
<td>2.7343</td>
<td>0.255</td>
</tr>
<tr>
<td>3</td>
<td>−0.042</td>
<td>3.15</td>
<td>0.369</td>
</tr>
<tr>
<td>4</td>
<td>0.132</td>
<td>7.3006</td>
<td>0.121</td>
</tr>
<tr>
<td>5</td>
<td>0.003</td>
<td>7.3031</td>
<td>0.199</td>
</tr>
<tr>
<td>6</td>
<td>−0.032</td>
<td>7.5482</td>
<td>0.273</td>
</tr>
</tbody>
</table>
Now let’s take the square of returns and repeat the exercise. The results are shown in Table 1.4. As can be seen from this table, there is a strong evidence of serial correlation in the squared returns. The implications of this will be studied in Unit 5.

### Table 1.4 Square of Monthly Returns, FTSE

<table>
<thead>
<tr>
<th>Lag</th>
<th>Autocorrelation</th>
<th>Ljung-Box Q Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.147</td>
<td>5.080</td>
<td>0.024</td>
</tr>
<tr>
<td>2</td>
<td>0.069</td>
<td>6.198</td>
<td>0.045</td>
</tr>
<tr>
<td>3</td>
<td>0.242</td>
<td>20.080</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.124</td>
<td>23.746</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.091</td>
<td>25.714</td>
<td>0.000</td>
</tr>
<tr>
<td>6</td>
<td>0.090</td>
<td>27.672</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### 1.6 Linear Dependency across Asset Returns

So far, we have focused on some of the stylised facts about individual series of asset returns. In this section, we shift the focus towards the dependence of returns across assets. As you may recall from previous courses, the widely used common measure of dependence is the correlation coefficient (also known as Pearson’s correlation), which is given by the following equation

\[
\rho[X,Y] = \frac{\text{Cov}[X,Y]}{\sqrt{\text{V}(X)\text{V}(Y)}}
\]  

(1.36)

where \( \text{Cov}(X,Y) \) is the covariance between \( X \) and \( Y \), \( \text{V}(X) \) is the variance of \( X \), and \( \text{V}(Y) \) is the variance of \( Y \). The correlation coefficient must lie between -1 and 1, with a zero value indicating no correlation between the two series. An estimator of the correlation coefficient is given by the following:

\[
\hat{\rho} = \frac{\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{T} (x_t - \bar{x})^2 \sum_{t=1}^{T} (y_t - \bar{y})^2}}
\]  

(1.37)

As you can see from Table 2.6 in your reading, the correlation between the various stock market indices is positive, implying that stock indices tend to move together. Another interesting observation is that the correlation tends to increase in turbulent times (for example, in times of crisis, the correlation between the indices becomes more positive). However, as discussed in your reading, this finding could be a spurious outcome and driven mainly by increased volatility.

What matters for us in this course is the possibility of jointly modelling asset returns and their volatility. In Unit 6, you are introduced to the multivariate GARCH models, which are an extension of the univariate GARCH models discussed in Unit 5. As you will see in Unit 6, multivariate GARCH models provide us with a useful tool to model time-varying autocorrelation. This
would allow us to identify whether there have been structural breaks in the correlation coefficient over time.

**Reading**

I would like you now to read Sections 2.4.1 and 2.4.2 of Jondeau, Poon and Rockinger.

**Example**

The data set contains weekly prices for FTSE, DAX, CAC and NYSE for the period January 1991 to August 2009. Table 1.5 provides summary statistics of the log weekly return for the various stock market indices. As can be seen from this table, the FTSE and NYSE show the highest kurtosis and skewness. As expected, we can reject the null hypothesis of normality for all of the indices.

**Table 1.5  Log Weekly Returns, NYSE, CAC, DAX and FTSE**

<table>
<thead>
<tr>
<th>Summary Statistics</th>
<th>RETURNNYSE</th>
<th>RETURNCAC</th>
<th>RETURNDAX</th>
<th>RETURNFTSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.001318</td>
<td>0.000873</td>
<td>0.001404</td>
<td>0.000849</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>0.003316</td>
<td>0.001612</td>
<td>0.003477</td>
<td>0.002114</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>0.121278</td>
<td>0.124321</td>
<td>0.149421</td>
<td>0.125832</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>-0.217345</td>
<td>-0.250504</td>
<td>-0.243470</td>
<td>-0.236316</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>0.023283</td>
<td>0.029747</td>
<td>0.031876</td>
<td>0.024017</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>-1.064534</td>
<td>-0.745220</td>
<td>-0.627666</td>
<td>-0.948353</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>14.02785</td>
<td>9.137679</td>
<td>8.457588</td>
<td>14.85545</td>
</tr>
<tr>
<td><strong>Jarque-Bera</strong></td>
<td>5108.93</td>
<td>1615.647</td>
<td>1270.126</td>
<td>5838.045</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>972</td>
<td>972</td>
<td>972</td>
<td>972</td>
</tr>
</tbody>
</table>

The correlation between the various weekly returns are summarised in Table 1.6.

**Table 1.6  Log Weekly Returns, Correlations**

<table>
<thead>
<tr>
<th>Column1</th>
<th>RETURNNYSE</th>
<th>RETURNFTSE</th>
<th>RETURNCAC</th>
<th>RETURNDAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>RETURNNYSE</td>
<td>1</td>
<td>0.194463414</td>
<td>0.209546676</td>
<td>0.191483563</td>
</tr>
<tr>
<td>RETURNFTSE</td>
<td>0.194463414</td>
<td>1</td>
<td>0.81628067</td>
<td>0.774448152</td>
</tr>
<tr>
<td>RETURNCAC</td>
<td>0.209546676</td>
<td>0.81628067</td>
<td>1</td>
<td>0.841403846</td>
</tr>
<tr>
<td>RETURNDAX</td>
<td>0.191483563</td>
<td>0.774448152</td>
<td>0.841403846</td>
<td>1</td>
</tr>
</tbody>
</table>

**Note:** Eviews automatically provides the correlation matrix. Once you have opened a file go to:

1. Quick
2. Scroll down to ‘Group Statistics’
3. Then choose ‘Correlations’

To provide a summary table for all of the series, choose
Quick
2 Scroll down to ‘Group Statistics’
3 Then choose ‘Descriptive Statistics’
4 Then choose ‘Individual samples’
This is a very useful tool when your dataset contains more than one series.

*Eviews* presents these correlations in a matrix form (matrix algebra is the subject of the next unit). But for now, notice that elements in the main diagonal all take the value of 1 since these measure the correlation of the returns of a certain index with itself. The off-diagonal elements measure the sample correlation across the various indices. Interestingly, the correlation matrix shows high correlation among the log weekly returns of the European stock indices. For instance, the correlation between the weekly log returns of FTSE (RETURNFTSE) and the weekly log returns of CAC (RETURNCAC) is 81%, and with weekly log return of DAX (RETURNDAX) is 77%, whereas the correlation with the weekly log return of NYSE (RETURNNYSE) is less than 20%. The highest correlation is between RETURNDAX and RETURNCAC.

As implied in your reading, it is highly unlikely for the correlation to remain constant throughout the entire sample. Thus, it is worth estimating the time varying correlation. This will be the subject of Unit 6. But just to anticipate the discussion of Unit 6, Figures 1.4 and 1.5 show the time varying correlation between RETURNFTSE and RETURNNYSE and between RETURNFTSE and RETURNDAX using a six-month rolling window.²

**Figure 1.4** Time-Varying Correlation between Weekly Log Returns of FTSE and NYSE

These have been calculated using Excel. It is quite clear that the correlation coefficient exhibits very volatile behaviour. Interestingly, while the correlation coefficient between RETURNFTSE and RETURNDAX is strong and has increased in the later part of the data period (Figure 1.5), the correlation between RETURNFTSE and RETURNNYSE is

---
² This involves calculating the correlation coefficients for the first six months, and then rolling the sample forward by including a new observation and dropping the first observation from the previous sample. It is an arbitrary method but an easy way to calculate the time varying correlation.
very volatile, taking negative and positive values, and evolving around zero in the later part of the data period (Figure 1.4). What is important to stress, though, is that the correlation is not constant and tends to vary over time.

Figure 1.5 Time-Varying Correlation between Weekly Log Returns for FTSE and DAX

1.7 Conclusion
The main objective of this unit has been to analyse some of the properties of financial asset returns. Please check that you have achieved the Learning Outcomes listed at the start of the unit. They are repeated here, so that you can now test yourself against them. You should now be able to

- define and compute the various measures of financial returns, including the simple return, gross return, multiperiod returns, continuously compounded returns
- calculate the sample moments of financial returns, including the skewness and kurtosis of financial returns, using Eviews;
- explain and discuss some of the stylised statistical properties of asset returns
- analyse and appreciate the issue of time dependency in asset returns;
- analyse the linear dependence across financial assets.

References


Exercises

1 Goto the website http://uk.finance.yahoo.com, and download daily data for the Dow Chemical Company stock over the years 2004 to 2008. Use the close price adjusted for dividends and splits. (Hint: For the Company or symbol, type ‘DOW’ and for the market, choose USA. Then choose Historical Prices. For reference, on 2 January 2004 the adjusted closing price was 32.52, on 5 January 2004, 33.14, and on 31 December 2008, 14.47.)

   a Using an Excel spread sheet (or similar package), calculate the daily simple return, and the daily log return for the stock.

   b Using the daily log returns, compute the continuously compounded annual return for 2004 to 2008. Plot the annual return on a graph and comment on the graph.

   c Transfer the data to Eviews. Calculate the sample mean, standard deviation, skewness and kurtosis of the daily log returns.

   d Are the daily log returns for Dow normally distributed? Explain how you can formally test for the null hypothesis of normality. What are the main limitations of the J-B test?

2 The text file c359_u1_q2.txt contains weekly data on the oil price and the share prices for ConocoPhillips for the period 2000 to 2008.

   a Calculate the sample mean, standard deviation, skewness and kurtosis of the weekly log returns for both the oil price and the share price.

   b Compute the correlation coefficient between the two series. Comment on the results.

3 The text file c359_u1_q3.txt contains weekly data on oil prices and the US trade-weighted exchange rate index against major currencies, from January 2000 to August 2009.

   a Calculate the sample mean, standard deviation, skewness and kurtosis of weekly log returns for both the oil price and the exchange rate.

   b Compute the correlation coefficient between the two series. Comment on the results.

   c Using Excel (or a similar package), compute the time varying correlation using a rolling six-month window. Alternatively, you can use the Eviews function @movcor(x,y,n) which provides an n-period moving correlation, calculating the correlation between x and y of the current and n-1 previous observations. NAs are propagated with @movcor, and NAs are not propagated with @mcor. Comment on the results.
Answers to Exercises

1 Dow Chemical Company (DOW) daily data:

a See the Excel file c359_u1_q1.xls (1997-2003 compatible) for the calculations. (If you have problems downloading the data, the tab-delimited text file c359_u1_q1.txt contains the date and daily adjusted closing price for Dow, 2004-2008.)

b See the Excel file c359_u1_q1.xls (1997-2003 compatible) for the calculations. The annualised compounded annual return (Figure 1.6) exhibits a pattern of cyclicality. The year 2008 witnessed a huge decline in the annualised returns as the financial crisis and the collapse in international trade had a negative impact on the petrochemical sector.

Figure 1.6 Annualised Compounded Annual Return, Dow Chemical Company

![Bar Chart](image)

The sample mean, standard deviation, skewness and kurtosis of daily log returns are shown in Figure 1.7. The sample mean is close to zero while the standard deviation is quite high. The maximum-minimum range is quite wide, ranging from 10% to -21%. The series shows evidence of excess kurtosis and some skewness.

c Based on Figure 1.7 below, it is clear that the daily log returns for Dow are not normally distributed. To formally test for normality, we use the Jarque-Bera test. As can be seen from this test, we can strongly reject the null hypothesis of normality.

d Based on Figure 1.7 below, it is clear that the daily log returns for Dow are not normally distributed. To formally test for normality, we use the Jarque-Bera test. As can be seen from this test, we can strongly reject the null hypothesis of normality.
Your reading mentions two main limitations of the Jarque-Bera test. The first is that the test is asymptotically distributed and the test is valid for large samples. This is not a major limitation in our case because our sample is quite large. The second limitation is that the empirical skewness and kurtosis are computed for given values of mean and variance. These two are subject to sampling errors. This is a common problem for all tests of normality.

2 Weekly data on the oil price and the share prices for ConocoPhillips:

a The sample mean, standard deviation, skewness and kurtosis are shown for log oil returns (Figure 1.8) and ConocoPhillips stock returns (Figure 1.9).

Like stock returns, oil returns exhibit strong kurtosis and negative skewness driven mainly by the collapse of the oil price in the second half of 2008. The Jarque-Bera test suggests that we can strongly the reject the null hypothesis of normality.
One would expect oil returns to be positively associated with an oil company’s stock returns. However, as seen in Table 1.7, this is not the case and the correlation is close to zero. This may suggest that getting exposure to the oil price by buying shares of oil-producing companies is not an effective strategy.

Table 1.7 Weekly Log Return, Oil Price and ConocoPhillips, Correlation

<table>
<thead>
<tr>
<th></th>
<th>CPRETURN</th>
<th>OILRETURN</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPRETURN</td>
<td>1</td>
<td>-0.02519</td>
</tr>
<tr>
<td>OILRETURN</td>
<td>-0.02519</td>
<td>1</td>
</tr>
</tbody>
</table>

Weekly data on oil prices and the US trade-weighted exchange rate index:

The sample mean, standard deviation, skewness and kurtosis are shown for log exchange rate returns (Figure 1.10) and log oil returns (Figure 1.11). Like stock returns and commodity returns, exchange rate returns exhibit excess kurtosis. The Jarque-Bera test suggests that we can strongly reject the null hypothesis of normality.
Table 1.8 indicates the correlation between the two series is negative i.e. a depreciation of the US currency against other major currencies is associated with positive oil returns.

**Table 1.8  Weekly Log Returns, Exchange Rate and Oil Price, Correlation**

<table>
<thead>
<tr>
<th></th>
<th>ERRETURN</th>
<th>OILRETURN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERRETURN</td>
<td>1</td>
<td>-0.13729</td>
</tr>
<tr>
<td>OILRETURN</td>
<td>-0.13729</td>
<td>1</td>
</tr>
</tbody>
</table>
The time varying correlation is shown in Figure 1.12. The time varying correlation shows quite clearly that the correlation is far from constant. For instance, between July and October 2008 the degree of (negative) correlation increased dramatically, to disappear again during the first five months of 2009, and to reappear again between May and August 2009. Given this behaviour, one may wonder whether the two series are correlated at all!

Figure 1.12 Time-Varying Correlation, Weekly Log Returns, Oil Price and Exchange Rate