14 Conducting empirical research or doing a project or dissertation in finance

Learning outcomes

In this chapter, you will learn how to

- Choose a suitable topic for an empirical research project in finance
- Draft a research proposal
- Find appropriate sources of literature and data
- Determine a sensible structure for the dissertation
- Set up and conduct a valid event study
- Employ the Fama–MacBeth and Fama–French approaches to testing asset pricing models and explaining the variation in asset returns

14.1 What is an empirical research project and what is it for?

Many courses, at both the undergraduate and postgraduate levels, require or allow the student to conduct a project. This may vary from being effectively an extended essay to a full-scale dissertation or thesis of 10,000 words or more.

Students often approach this part of their degree with much trepidation, although in fact doing a project gives students a unique opportunity to select a topic of interest and to specify the whole project themselves from start to finish. The purpose of a project is usually to determine whether students can define and execute a piece of fairly original research within given time, resource and report-length constraints. In terms of econometrics, conducting empirical research is one of the best ways to get to grips with the theoretical material, and to find out what practical difficulties econometricians encounter when conducting research. Conducting the research gives the investigator the opportunity to solve a puzzle and potentially to uncover something that nobody else has; it can be a highly rewarding experience. In addition, the project allows students to select a topic of direct interest or relevance to them, and is often useful in helping students to develop time-management and report-writing skills. The final document can in many cases provide a platform for discussion at job interviews, or act as a springboard to further study at the taught postgraduate or doctoral level.

This chapter seeks to give suggestions on how to go about the process of conducting empirical research in finance. Only general guidance is given, and following this advice cannot necessarily guarantee high marks, for the objectives and required level of the project will vary from one institution to another.[1]

14.2 Selecting the topic

Following the decision or requirement to do a project, the first stage is to determine an appropriate *subject area*. This is, in many respects, one of the most difficult and most crucial parts of the whole exercise. Some students are immediately able to think of a precise topic, but for most, it is a process that starts with specifying a very general and very broad subject area, and subsequently narrowing it down to a much smaller and manageable problem.

Inspiration for the choice of topic may come from a number of sources. A good approach is to think rationally about your own interests and areas of expertise. For example, you may have worked in the financial markets in some capacity, or you may have been particularly interested in one aspect of a course unit that you have studied. It is worth spending time talking to some of your instructors in order to gain their advice on what are interesting and plausible topics in their subject areas. At the same time, you may feel very confident at the quantitative end of finance, pricing assets or estimating models for example, but you may not feel comfortable with qualitative analysis where you are asked to give an opinion on particular issues (e.g. 'should financial markets be more regulated?'). In that case, a highly technical piece of work may be appropriate.

Equally, many students find econometrics both difficult and uninteresting. Such students may be better suited to more qualitative topics, or topics that involve only elementary statistics, but where the rigour and value added comes from some other aspect of the problem. A case-study approach that is not based on any quantitative analysis may be entirely acceptable and indeed an examination of a set of carefully selected case studies may be more appropriate for addressing particular problems, especially in situations where hard data are not readily available, or where each entity is distinct so that generalising from a model estimated on one set of data may be inadvisable. Case studies are useful when the case itself is unusual or unique or when each entity under study is very heterogeneous. They involve more depth of study than quantitative approaches. Highly mathematical work that has little relevance and which has been applied inappropriately may be much weaker than a well constructed and carefully analysed case study.

Combining all of these inputs to the choice of topic should enable you at the least to determine whether to conduct quantitative or non-quantitative work, and to select a general subject area (e.g. pricing securities, market microstructure, risk management, asset selection, operational issues, international finance, financial

[1] Note that there is only one review question for this chapter and that is to write an excellent research project.

Box 14.1 Possible types of research project

- An empirical piece of work involving quantitative analysis of data
- A survey of business practice in the context of a financial firm
- A new method for pricing a security, or the theoretical development of a new method for hedging an exposure
- A critical review of an area of literature
- An analysis of a new market or new asset class.

Each of these types of project requires a slightly different approach, and is conducted with varying degrees of success. The remainder of this chapter focuses upon the type of study which involves the formulation of an empirical model using the tools developed in this book. This type of project seems to be the one most commonly selected. It also seems to be a lower risk strategy than others. For example, projects which have the bold ambition to develop a new financial theory, or a whole new model for pricing options, are likely to be unsuccessful and to leave the student with little to write about. Also, critical reviews often lack rigour and are not critical enough, so that an empirical application involving estimating an econometric model appears to be a less risky approach, since the results can be written up whether they are 'good' or not.

econometrics, etc.). The project may take one of a number of forms as illustrated in box 14.1.

A good project or dissertation must have an element of *originality*, i.e. a 'contribution to knowledge'. It should add, probably a very small piece, to the overall picture in that subject area, so that the body of knowledge is larger at the end than before the project was started. This statement often scares students, for they are unsure from where the originality will arise. In empirically based projects, this usually arises naturally. For example, a project may employ standard techniques on data from a different country or a new market or asset, or a project may develop a new technique or apply an existing technique to a different area. Interesting projects can often arise when ideas are taken from another field and applied to finance – for example, you may be able to identify ideas or approaches from the material that you studied from a different discipline as part of your undergraduate degree.

A good project will also contain an in-depth analysis of the issues at hand, rather than a superficial, purely descriptive presentation, as well as an individual contribution. A good project will be interesting, and it will have relevance for one or more user groups (although the user group may be other academic researchers and not necessarily practitioners); it may or may not be on a currently fashionable and newsworthy topic. The best research challenges prior beliefs and changes the way that the reader thinks about the problem under investigation. Good projects can be primarily of interest to other academics and they do not necessarily have to be of direct practical applicability. On the other hand, highly practical work must also be well grounded in the academic approach to doing research.

The next stage is to transform this broad direction into a workably sized topic that can be tackled within the constraints laid down by the institution. It is important to ensure that the aims of the research are not so broad or substantive that the questions cannot be addressed within the constraints on available time and word limits. The objective of the project is usually not to solve the entire world's financial puzzles, but rather to form and address a small problem.

It is often advisable at this stage to browse through recent issues of the main journals relevant to the subject area. This will show which ideas are relatively fashionable, and how existing research has tackled particular problems. A list of relevant journals is presented in table 14.1. They can be broadly divided into two categories: practitioner-oriented and academic journals. Practitioner-oriented journals are usually very focused in a particular area, and articles in these often centre on very practical problems, and are typically less mathematical in nature and less theory-based, than are those in academic journals. Of course, the divide between practitioner and academic journals is not a total one, for many articles in practitioner journals are written by academics and vice versa! The list given in table 14.1 is by no means exhaustive and, particularly in finance, new journals appear on a monthly basis.

Many web sites contain lists of journals in finance or links to finance journals. Some useful ones are:

- www.cob.ohio-state.edu/dept/fin/overview.htm – the Virtual Finance Library, with good links and a list of finance journals
- www.helsinki.fi/WebEc/journals.html – provides a list of journals in the economics area, including finance, plus a number of finance-related resources
- www.people.hbs.edu/pgompers/finjourn.htm – provides a list of links to finance journals
- www.numa.com/ref/journals.htm – the Numa directory of derivatives journals – lots of useful links and contacts for academic and especially practitioner journals on derivatives
- www.aeaweb.org/econlit/journal_list.php – provides a comprehensive list of journals in the economics area, including finance

14.3 Sponsored or independent research?

Some business schools are sufficiently well connected with industry that they are able to offer students the opportunity to work on a specific research project with a 'sponsor'. The sponsor may choose the topic and offer additional expert guidance from a practical perspective. Sponsorship may give the student an insight into the kind of research problems that are of interest to practitioners, and will probably ensure that the work is practically focused and of direct relevance in the private sector. The sponsor may be able to provide access to proprietary or confidential data, which will broaden the range of topics that could be tackled. Most importantly, many students hope that if they impress the firm that they are working with, a permanent job offer will follow.

Table 14.1 Journals in finance and econometrics

Journals in finance	Journals in econometrics and related
Applied Financial Economics	Biometrika
Applied Mathematical Finance	Econometrica
European Financial Management	Econometric Reviews
European Journal of Finance	Econometric Theory
Finance and Stochastics	Econometrics Journal
Financial Analysts Journal	International Journal of Forecasting
Financial Management	Journal of Applied Econometrics
Financial Review	Journal of Business and Economic Statistics
Global Finance Journal	Journal of Econometrics
International Journal of Finance & Economics	Journal of Forecasting
International Journal of Theoretical and Applied Finance	Journal of the American Statistical Association
Journal of Applied Corporate Finance	Journal of Financial Econometrics
International Review of Financial Analysis	Journal of the Royal Statistical Society (A to C)
Journal of Applied Finance	Journal of Time Series Analysis
Journal of Asset Management	Society for Nonlinear Dynamics and Econometrics
Journal of Banking and Finance	
Journal of Business	
Journal of Business Finance & Accounting	
Journal of Computational Finance	
Journal of Corporate Finance	
Journal of Derivatives	
Journal of Empirical Finance	
Journal of Finance	
Journal of Financial & Quantitative Analysis	
Journal of Financial Economics	
Journal of Financial Markets	
Journal of Financial Research	
Journal of Fixed Income	
Journal of Futures Markets	
Journal of International Financial Markets, Institutions and Money	
Journal of International Money and Finance	
Journal of Money, Credit, and Banking	
Journal of Portfolio Management	
Journal of Risk	
Journal of Risk and Insurance	
Journal of Risk and Uncertainty	
Mathematical Finance	
Pacific Basin Finance Journal	
Quarterly Review of Economics and Finance	
Review of Asset Pricing Studies	
Review of Behavioural Finance	
Review of Corporate Finance Studies	
Review of Finance	
Review of Financial Studies	
Risk	

The chance to work on a sponsored project is usually much sought after by students but it is very much a double-edged sword, so that there are also a number of disadvantages. First, most schools are not able to offer such sponsorships, and even those that can are usually able to provide them to only a fraction of the class. Second, the disappointing reality is that the problems of most interest and relevance to practitioners are often (although admittedly not always) of less interest to an academic audience – fundamentally, the objectives of the sponsor and of a university may be divergent. For example, a stereotypical investment bank might like to see a project that compares a number of technical trading rules and evaluates their profitability; but many academics would argue that this area has been well researched before and that finding a highly profitable rule does not constitute a contribution to knowledge and is therefore weak as a research project. So if you have the opportunity to undertake a sponsored project, ensure that your research is of academic as well as practical value – after all, it will almost certainly be the academic who grades the work.

14.4 The research proposal

Some schools will require the submission of a research proposal which will be evaluated and used to determine the appropriateness of the ideas and to select a suitable supervisor. While the requirements for the proposal are likely to differ widely from one institution to another, there are some general points that may be universally useful. In some ways, the proposal should be structured as a miniature version of the final report, but without the results or conclusions!

- The required length of the proposal will vary, but will usually be between one and six sides of A4, typed with page numbering.
- The proposal should start by briefly motivating the topic – why is it interesting or useful?
- There should be a **brief** review of the relevant literature, but this should not cover more than around a third to one half of the total length of the proposal.
- The research questions or hypotheses to be tested should then be clearly stated.
- There should be a discussion of the data and methodology that you intend to use.
- Some proposals also include a time-scale – i.e. which parts of the project do you expect to have completed by what dates?

14.5 Working papers and literature on the internet

Unfortunately, the lag between a paper being written and it actually being published in a journal is often two–three years (and increasing fast), so that research in even the most recent issues of the published journals will be somewhat dated. Additionally, many securities firms, banks and central banks across the world, produce high quality research output in report form, which they often do not bother to try to publish. Much of this is now available on the internet, so it is worth conducting searches with keywords using readily available web search engines. A few suggestions for places to start are given in table 14.2.

Table 14.2 Useful internet sites for financial literature

Universities

Almost all universities around the world now make copies of their discussion papers available electronically.

A few examples from finance departments are:

http://w4.stern.nyu.edu/finance – Department of Finance, Stern School, New York University

http://fic.wharton.upenn.edu/fic/papers.html – Wharton Financial Institutions Center

http://haas.berkeley.edu/finance/WP/rpf.html – University of California at Berkeley

www.icmacentre.ac.uk/research/discussion-papers – ICMA Centre, University of Reading, of course!

US Federal Reserve Banks and the Bank of England

www.bankofengland.co.uk – Bank of England – containing their working papers, news and discussion

www.frbatlanta.org – Federal Reserve Bank of Atlanta – including information on economic and research data and publications

www.stls.frb.org/fred – Federal Reserve Bank of St. Louis – a great deal of useful US data, including monetary, interest rate, and financial data, available daily, weekly, or monthly, including long time histories of data

www.chicagofed.org – Federal Reserve Bank of Chicago – including interest data and useful links

www.dallasfed.org – Federal Reserve Bank of Dallas – including macroeconomic, interest rate, monetary and bank data

www.federalreserve.gov/pubs/ifdp – Federal Reserve Board of Governors International Finance Discussion Papers

www.ny.frb.org/research – Federal Reserve Bank of New York

International bodies

http://dsbb.imf.org – the International Monetary Fund (IMF) – including working papers, forecasts, and IMF primary commodity price series

www.worldbank.org/reference – World Bank working papers in finance

www.oecd-ilibrary.org – Organisation for Economic Cooperation and Development (OECD) working papers, data etc., searchable

Table 14.2 (*cont.*)

Miscellaneous

www.nber.org – National Bureau of Economic Research (NBER) – huge database of discussion papers and links including data sources

http://econpapers.repec.org – Econpapers (formerly WoPEc) – huge database of working papers in areas of economics, including finance

www.ssrn.com – The Social Science Research Network – a huge and rapidly growing searchable database of working papers and the abstracts of published papers

The free data sources used in this book

www.nationwide.co.uk/default.htm – UK house price index, quarterly back to 1952, plus house prices by region and by property type

www.oanda.com/convert/fxhistory – historical exchange rate series for an incredible range of currency pairs

www.bls.gov – US Bureau of Labor Statistics – US macroeconomic series

www.federalreserve.gov/econresdata/default.htm – US Federal Reserve Board – more US macroeconomic series, interest rates, etc. and working papers

http://research.stlouisfed.org/fred2 – a vast array of US macroeconomic series

http://finance.yahoo.com – Yahoo! Finance – an incredible range of free financial data, information, research and commentary

14.6 Getting the data

Although there is more work to be done before the data are analysed, it is important to think prior to doing anything further about *what data are required* to complete the project. Many interesting and sensible ideas for projects fall flat owing to a lack of availability of relevant data. For example, the data required may be confidential, they may be available only at great financial cost, they may be too time-consuming to collect from a number of different paper sources, and so on. So before finally deciding on a particular topic, make sure that the data are going to be available.

The data may be available at your institution, either in paper form (for example, from the IMF or World Bank reports), or preferably electronically. Many universities have access to Reuters, Datastream or the Bloomberg. Many of the URLs listed above include extensive databases and furthermore, many markets and exchanges have their own web pages detailing data availability. One needs to be slightly careful, however, in ensuring the accuracy of freely available data; 'free' data also sometimes turn out not to be!

14.7 Choice of computer software

Clearly, the choice of computer software will depend on the tasks at hand. Projects that seek to offer opinions, to synthesise the literature and to provide a review, may not require any specialist software at all. However, even for those conducting highly technical research, project students rarely have the time to learn a completely new programming language from scratch while conducting the research. Therefore, it is usually advisable, if possible, to use a standard software package. It is also worth stating that marks will hardly ever be awarded for students who 'reinvent the wheel'. Therefore, learning to program a multivariate GARCH model estimation routine in C++ may be a valuable exercise for career development for those who wish to be quantitative researchers, but is unlikely to attract high marks as part of a research project unless there is some other value added. The best approach is usually to conduct the estimation as quickly and accurately as possible to leave time free for other parts of the work.

14.8 Methodology

Good research is rarely purely empirical – the empirical model should arise from an economic or financial *theory* and this theory should be presented and discussed before the investigative work begins. We could define a theory as a system of statements that encompass a number of hypotheses. Theory shows what features in the data and what relationships would be expected based on some underlying principles. Theory can give order and meaning to empirical results, and can ensure that the findings are not the result of a data-mining exercise.

Assuming that the project is empirical in nature (i.e. it seeks to test a theory or answer a particular question using actual data), then an important question will concern the type of model to employ. This chapter will now discuss two of the most important approaches to conducting research in finance that have emerged over the past two or three decades: the event study methodology and the Fama–French approach. Although neither of these requires any new econometric tools that were not covered in previous chapters, the terminology employed is quite specific to this part of the literature and thus a focused discussion of how to implement these techniques may prove useful.

14.9 Event studies

Event studies are very useful in finance research and as a result they are extremely commonly employed in the literature. In essence, they represent an attempt to gauge the effect of an identifiable *event* on a financial variable, usually stock returns. So, for example, research has investigated the impact of various types of announcements (e.g. dividends, stock splits, entry into or deletion from a stock index) on the returns of the stocks concerned. Event studies are often considered to be tests for market efficiency: if the financial markets are informationally efficient, there should be an immediate reaction to the event on the announcement date and no further reaction on subsequent trading days.

MacKinlay (1997) argues that conducting event studies initially appears difficult but is in fact easy; my view is that exactly the reverse is true: in principle, event studies are simple to understand and to conduct, but to do so in a rigorous manner requires a great deal of thought. There is a bewildering array of approaches that can be deployed, and at first blush it is not at all clear which of them is appropriate or optimal. The main groundwork for conducting modern event studies was established by Ball and Brown (1968) and by Fama *et al.* (1969), but as MacKinlay notes, something like them was conducted more than three decades earlier.

While there are now many useful survey papers that describe the various aspects of event studies in considerable detail, unfortunately each has its own notation and approach which can be confusing. Corrado (2011) is a recent example, although Armitage (1995) and MacKinlay (1997) are particularly clearly explained and more closely resemble the treatment given here. A similar discussion is offered by Campbell *et al.* (1997) but using matrix notation.

14.9.1 Some notation and a description of the basic approach

We of course need to be able to define precisely the dates on which the events occur, and the sample data are usually aligned with respect to this date. If we have N events in the sample, we usually specify an 'event window', which is the period of time over which we investigate the impact of the event. The length of this window will be set depending on whether we wish to investigate the short- or long-run effects of the event. It is common to examine a period comprising, say, ten trading days before the event up to ten trading days after as a short-run event window, while long-run windows can cover a month, a year, or even several years after the event.

A first question to ask once the event has been identified is what frequency of data should be employed in the analysis. MacKinlay (1997) shows that the power of event studies to detect abnormal performance is much greater when daily data are employed rather than weekly or monthly observations, so that the same power can be achieved with a much smaller N, or for given N, the power will be much larger. Although it would in some cases be possible to use intra-daily data, these are harder to collect and may bring additional problems including nuisance microstructure effects; this is perhaps why daily observations are the frequency of choice for most studies in the literature.[2]

Define the return for each firm i on each day t during the event window as R_{it}. We can conduct the following approach separately for each day within the event window – for example, we might investigate it for all of ten days before the event up to ten days after (where $t = 0$ represents the date of the event and $t = -10, -9, -8, \ldots, -1, 0, 1, 2, \ldots, 8, 9, 10$). Note that we will need to

[2] We need to be aware of the potential impacts that thin trading of stocks may have, leading to stale prices and unrepresentative abnormal returns; however, this issue is not discussed further here.

exercise care in the definition of the reference day $t = 0$ if the announcement is made after the close of the market.

In most cases, we need to be able to separate the impact of the event from other, unrelated movements in prices. For example, if it is announced that a firm will become a member of a widely followed stock index and its share price that day rises by 4%, but on average the prices of all other stocks also rise by 4%, it would be unwise to conclude that all of the increase in the price of the stock under study is attributable to the announcement. This motivates the idea of constructing abnormal returns, denoted AR_{it}, which are calculated by subtracting an expected return from the actual return

$$AR_{it} = R_{it} - E(R_{it}) \tag{14.1}$$

There are numerous ways that the expected returns can be calculated, but usually this is achieved using a sample of data before the event window so that the nature of the event is not allowed to 'contaminate' estimation of the expected returns. Armitage (1995) suggests that estimation periods can comprise anything from 100 to 300 days for daily observations and 24 to 60 months when the analysis is conducted on a monthly basis. Longer estimation windows will in general increase the precision of parameter estimation, although with it the likelihood of a structural break and so there is a trade-off.

If the event window is very short (e.g. a day or a few days), then we need be far less concerned about constructing an expected return since it is likely to be very close to zero over such a short horizon. In such circumstances, it will probably be acceptable to simply use the actual returns in place of abnormal returns.

It is further often the case that a gap is left between the estimation period and the event window, to be completely sure that anticipation (i.e. 'leakage') of the event does not affect estimation of the expected return equation. However, it might well be the case that in practice we do not have the luxury of doing this since the sample period available is insufficient. Clearly, what we would like to do is to calculate the return that would have been expected for that stock if the event did not happen at all so that we can isolate the impact of the event from any unrelated incidents that may be occurring at the same time.

The simplest method for constructing expected returns (apart from setting them to zero) is to assume a constant mean return, so that the expected return is simply the average return for each stock i calculated at the same frequency over the estimation window, which we might term $\bar{R}_i$. Brown and Warner (1980, 1985) conduct a simulation experiment to compare methods of estimating expected returns for event studies. They find that an approach simply using historical return averages outperforms many more complicated approaches because of the estimation error that comes with the latter.

A second, slightly more sophisticated approach, is to subtract the return on a proxy for the market portfolio that day t from the individual return. This will certainly overcome the impact of general market movements in a rudimentary way, and is equivalent to the assumption that the stock's beta in the market model or the CAPM is unity.

Probably the most common approach to constructing expected returns, however, is to use the market model. This in essence works by constructing the expected return using a regression of the return to stock i on a constant and the return to the market portfolio

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it} \tag{14.2}$$

The expected return for firm i on any day t during the event window would then be calculated as the beta estimate from this regression multiplied by the actual market return on day t.

An interesting question is whether the expected return should incorporate the α from the estimation period in addition to β multiplied by the market return. Most applications of event studies include this, and indeed the original study by Fama *et al.* includes an alpha. However, we need to exercise caution when doing so since if – either because of some unrelated incident affecting the price of the stock or in anticipation of the event – the alpha is particularly high (particularly low) during the estimation period, it will push up (down) the expected return. Thus it may be preferable to assume an expected value of zero for the alpha and to exclude it from the event period abnormal return calculation.

In most applications, a broad stock index such as the FTSE All-Share or the S&P500 would be employed to proxy for the market portfolio. This equation can be made as complicated as desired – for example, by allowing for firm size or other characteristics – these would be included as additional factors in the regression with the expected return during the event window being calculated in a similar fashion. An approach based on the arbitrage pricing models of Chen *et al.* (1986) or of Fama and French (1993) could also be employed – more discussion is made of this issue in the following section.

A final further approach would be to set up a 'control portfolio' of firms that have characteristics as close as possible to those of the event firm – for example, matching on firm size, beta, industry, book-to-market ratio, etc. – and then using the returns on this portfolio as the expected returns. Armitage (1995) reports the results of several Monte Carlo simulations that compare the results of various model frameworks that can be used for event studies.

The hypothesis testing framework is usually set up so that the null to be examined is of the event having no effect on the stock price (i.e. an abnormal return of zero). Under the null of no abnormal performance for firm i on day t during the event window, we can construct test statistics based on the standardised abnormal performance. These test statistics will be asymptotically normally distributed (as the length of the estimation window, T, increases)

$$AR_{it} \sim N(0, \sigma^2(AR_{it}))$$

where $\sigma^2(AR_{it})$ is the variance of the abnormal returns, which can be estimated in various ways. A simple method, used by Brown and Warner (1980) among others, is to use the time series of data from the estimation of the expected returns separately for each stock. So we could define $\hat{\sigma}^2(AR_{it})$ as being the equal to the variance of the residuals from the market model, which could be calculated for example using

$$\hat{\sigma}^2(AR_{it}) = \frac{1}{T-2}\sum_{t=2}^{T}\hat{u}_{it}^2 \tag{14.3}$$

where T is the number of observations in the estimation period. If instead the expected returns had been estimated using historical average returns, we would simply use the variance of those.

Sometimes, an adjustment is made to $\hat{\sigma}^2(AR_{it})$ that reflects the errors arising from estimation of α and β in the market model. Including the adjustment, the variance in the previous equation becomes

$$\hat{\sigma}^2(AR_{it}) = \frac{1}{T-2}\sum_{t=2}^{T}\left(\hat{u}_{it}^2 + \frac{1}{T}\left[1 + \frac{R_{mt} - \bar{R}_m}{\hat{\sigma}_m^2}\right]\right) \tag{14.4}$$

where $\bar{R}_m$ and $\hat{\sigma}_m^2$ are the average and variance of the returns on the market portfolio respectively during the estimation window. It should be clear that as the length of the estimation period, T, increases, this adjustment will gradually shrink to zero.

We can then construct a test statistic by taking the abnormal return and dividing it by its corresponding standard error, which will asymptotically follow a standard normal distribution[3]

$$S\hat{A}R_{it} = \frac{\hat{A}R_{it}}{[\hat{\sigma}^2(AR_{it})]^{1/2}} \sim N(0,1) \tag{14.5}$$

where $S\hat{A}R_{it}$ stands for the standardised abnormal return, which is the test statistic for each firm i and for each event day t.

It is likely that there will be quite a bit of variation of the returns across the days within the event window, with the price rising on some days and then falling on others. As such, it would be hard to identify the overall patterns. We may therefore consider computing the time series cumulative average return over a multi-period event window (for example, over ten trading days) by summing the average returns over several periods, say from time T_1 to T_2

$$C\hat{A}R_i(T_1, T_2) = \sum_{t=T_1}^{T_2}\hat{A}R_{it} \tag{14.6}$$

Note that the time from T_1 to T_2 may constitute the entire event window or it might just be a sub-set of it. The variance of this $C\hat{A}R$ will be given by the number

[3] Note that in some studies, since the sample variance has to be estimated, the test statistic is assumed to follow a student's t distribution with $T-k$ degrees of freedom in finite samples, where k is the number of parameters estimated in constructing the measure of expected returns ($k=2$ for the market model). Provided that the estimation window is of a reasonable length (e.g. six months of trading days or more), it will be inconsequential whether the t or normal distributions are employed.

of observations in the event window plus one multiplied by the daily abnormal return variance calculated in equation (14.4) above

$$\hat{\sigma}^2(CAR_i(T_1, T_2)) = (T_2 - T_1 + 1)\hat{\sigma}^2(\hat{AR}_{it}) \quad (14.7)$$

This expression is essentially the sum of the individual daily variances over the days in T_1 to T_2 inclusive.[4]

We can now construct a test statistic for the cumulative abnormal return in the same way as we did for the individual dates, which will again be standard normally distributed

$$S\hat{CAR}_i(T_1, T_2) = \frac{\hat{CAR}_i(T_1, T_2)}{[\hat{\sigma}^2(CAR_i(T_1, T_2))]^{1/2}} \sim N(0, 1) \quad (14.8)$$

It is common to examine a pre-event window (to consider whether there is any anticipation of the event) and a post-event window – in other words, we sum the daily returns for a given firm i for days $t - 10$ to $t - 1$, say, and separately from $t + 1$ to $t + 10$, with the actual day of the event, t, being considered on its own.

Typically, some of the firms will show a negative abnormal return around the event when a positive figure was expected, and this is probably not very useful. But if we have N firms or N events, it is usually of more interest whether the return averaged across all firms is statistically different from zero than whether this is the case for any specific individual firm. We could define this average across firms for each separate day t during the event window as

$$\hat{AR}_t = \frac{1}{N}\sum_{i=1}^{N} \hat{AR}_{it} \quad (14.9)$$

This firm-average abnormal return, $\hat{AR}_t$ will have variance given by $1/N$ multiplied by the average of the variances of the individual firm returns

$$\hat{\sigma}^2(AR_t) = \frac{1}{N^2}\sum_{i=1}^{N} \hat{\sigma}^2(AR_{it}) \quad (14.10)$$

Thus the test statistic (the standardised return) for testing the null hypothesis that the average (across the N firms) return on day t is zero will be given by

$$S\hat{AR}_t = \frac{\hat{AR}_t}{[\hat{\sigma}^2(AR_t)]^{1/2}} = \frac{\frac{1}{N}\sum_{i=1}^{N} \hat{AR}_{it}}{[\frac{1}{N^2}\sum_{i=1}^{N} \hat{\sigma}^2(AR_{it})]^{1/2}} \sim N(0, 1) \quad (14.11)$$

Finally, we can aggregate both across firms and over time to form a single test statistic for examining the null hypothesis that the average multi-horizon (i.e. cumulative) return across all firms is zero. We would get an equivalent statistic whether we first aggregated over time and then across firms or the other way

[4] The number of days during the period T_1 to T_2 including both the end points is $T_2 - T_1 + 1$.

around. The CAR calculated by averaging across firms first and then cumulating over time could be written

$$\hat{CAR}(T_1, T_2) = \sum_{t=T_1}^{T_2} \hat{AR}_t \tag{14.12}$$

Or equivalently, if we started with the $CAR_i(T_1, T_2)$ separately for each firm, we would take the average of these over the N firms

$$\hat{CAR}(T_1, T_2) = \frac{1}{N}\sum_{i=1}^{N} \hat{CAR}_i(T_1, T_2) \tag{14.13}$$

To obtain the variance of this $\hat{CAR}(T_1, T_2)$ we could take $1/N$ multiplied by the average of the variances of the individual $\hat{CAR}_i$.

$$\hat{\sigma}^2(CAR(T_1, T_2)) = \frac{1}{N^2}\sum_{i=1}^{N} \hat{\sigma}^2(CAR_i(T_1, T_2)) \tag{14.14}$$

And again we can construct a standard normally distributed test statistic as

$$S\hat{CAR}(T_1, T_2) = \frac{\hat{CAR}(T_1, T_2)}{[\hat{\sigma}^2(CAR(T_1, T_2))]^{1/2}} \sim N(0, 1) \tag{14.15}$$

14.9.2 Cross-sectional regressions

The methodologies and formulae presented above provide various tools for examining whether abnormal returns are statistically significant or not. However, it will often be the case that we are interested in allowing for differences in the characteristics of a sub-section of the events and also examining the link between the characteristics and the magnitude of the abnormal returns. For example, does the event have a bigger impact on small firms? Or on firms which are heavily traded etc.? The simplest way to achieve this would be to calculate the abnormal returns as desired using something like equation (14.2) above and then to use these as the dependent variable in a cross-sectional regression of the form

$$AR_i = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \ldots + \gamma_M x_{Mi} + w_i \tag{14.16}$$

where AR_i is the abnormal return for firm i measured over some horizon, and x_{ji}, $(j = 1, \ldots, M)$ are a set of M characteristics that are thought to influence the abnormal returns, γ_j measures the impact of the corresponding variable j on the abnormal return, and w_i is an error term. We can examine the sign, size and statistical significance of γ_0 as a test for whether the average abnormal return is significantly different from zero after allowing for the impacts of the M characteristics. MacKinlay (1997) advocates the use of heteroscedasticity-robust standard errors in the regression.

The abnormal return used in this equation would typically be measured over several days (or perhaps even the whole event window), but it could also be based on a single day.

14.9.3 Complications when conducting event studies and their resolution

The above discussion presents the standard methodology that is commonly employed when conducting event studies, and most of the time it will provide appropriate inferences. However, as always in econometrics, the use of test statistics requires a number of assumptions about the nature of the data and models employed. Some of these assumptions will now be highlighted and their implications explored.

Cross-sectional dependence

A key assumption when the returns are aggregated across firms is that the events are independent of one another. Often, this will not be the case, particularly when the events are clustered through time. For example, if we were investigating the impact of index recompositions on the prices of the stocks concerned, these index constituents generally only change at specific times of the year. So, typically, a bunch of stocks will enter into an index on the same day, and then there may be no further such events for three or six months.

The impact of this clustering is that we cannot assume the returns to be independent across firms, and as a result the variances in the aggregates across firms (equations (14.10) and (14.14)) will not apply since these derivations have effectively assumed the returns to be independent across firms so that all of the covariances between returns across firms can be set to zero. An obvious solution to this would be not to aggregate the returns across firms, but simply to construct the test statistics on an event-by-event basis and then to undertake a summary analysis of them (e.g. reporting their means, variances, percentage of significant events, etc.).

A second solution would be to construct portfolios of firms having the event at the same time and then the analysis would be done on each of the portfolios. The standard deviation would be calculated using the cross-section of those portfolios' returns on day t (or on days T_1 to T_2, as desired). This approach will allow for cross-correlations since they will automatically be taken into account in constructing the portfolio returns and the standard deviations of those returns. But a disadvantage of this technique is that it cannot allow for different variances for each firm as all are equally weighted within the portfolio; the standard method described above would do so, however.

Changing variances of returns

It has been argued in the literature that often the variance of returns will increase over the event window, but the variance figure used in the testing procedure will have been calculated based on the estimation window, which is usually some time before the event. Either the event itself or the factors that led to it are likely to increase uncertainty and with it the volatility of returns. As a result, the measured variance will be too low and the null hypothesis of no abnormal return during the event will be rejected too often. To deal with this, Boehmer *et al.* (1991), among others, suggest estimating the variance of abnormal returns by employing the cross-sectional variance of returns across firms during the *event* window. Clearly, if we adopt this procedure we cannot estimate separate test statistics for each firm (although arguably these are usually of little interest anyway). The variance estimator in equation (14.10) would be replaced by

$$\hat{\sigma}^2(AR_t) = \frac{1}{N^2}\sum_{i=1}^{N}(\hat{AR}_{it} - \hat{AR}_t)^2 \tag{14.17}$$

with the test statistic following as before. A similar adjustment could be made for the variance of the cumulative abnormal return

$$\sigma^2(CAR(T_1, T_2)) = \frac{1}{N^2}\sum_{i=1}^{N}\left(C\hat{AR}_i(T_1, T_2) - C\hat{AR}(T_1, T_2)\right) \tag{14.18}$$

While this test statistic will allow for the variance to change over time, a drawback is that it does not allow for differences in return variances across firms and nor does it allow for cross-correlations in returns caused by event clustering.

Weighting the stocks

Another issue is that the approach as stated above will not give equal weight to each stock's return in the calculation. The steps outlined above construct the cross-firm aggregate return (in equation (14.9)) and then standardise this using the aggregate standard deviation (in equation (14.11)). An alternative method would be to first standardise each firm's abnormal return (dividing by its appropriate standard deviation) and then aggregating these standardised abnormal returns. If we take the standardised abnormal return for each firm, $S\hat{AR}_{i,t}$, from equation (14.5), we can calculate the average of these across the N firms

$$S\hat{AR}_t = \frac{1}{N}\sum_{i=1}^{N} S\hat{AR}_{it} \tag{14.19}$$

These SARs have already been standardised so there is no need to divide them by the square root of the variance. If we take this SAR_t and multiply it by $\sqrt{N}$, we will get a test statistic that is asymptotically normally distributed and which, by construction, will give equal weight to each SAR (because we have taken an unweighted average of them)

$$\sqrt{N}SAR_t \sim N(0, 1)$$

We could similarly take an unweighted average of the standardised cumulative abnormal returns ($SCAR$)

$$SC\hat{AR}(T_1, T_2) = \frac{1}{N}\sum_{i=1}^{N} SC\hat{AR}_i(T_1, T_2) \tag{14.20}$$

and

$$\sqrt{N}SCAR(T_1, T_2) \sim N(0, 1)$$

If the true abnormal return is similar across securities, we would be better to equally weight the abnormal returns in calculating the test statistics (as in equations (14.19) and (14.20)), but if the abnormal return varies positively with its variance measure, then it would be better to give more weight to stocks with lower return variances (as in equation (14.15) for example).

Long event windows

Event studies are joint tests of whether the event-induced abnormal return is zero and whether the model employed to construct expected returns is correct. If we wish to examine the impact of an event over a long period (say, more than a few months), we need to be more careful about the design of the model for expected returns and also to ensure that this model appropriately allows for risk. Over short windows, discrepancies between models are usually small and any errors in model specification are almost negligible. But over the longer run, small errors in setting up the asset pricing model can lead to large errors in the calculation of abnormal returns and therefore the impact of the event.

A key question in conducting event studies to measure long-run impacts is whether to use cumulative abnormal returns (CARs), as described above, or buy-and-hold abnormal returns (BHARs). There are several important differences between the two. First, BHARs employ geometric returns rather than arithmetic returns (used in computing CARs) in calculating the overall return over the event period of interest. Thus the BHAR can allow for compounding whereas the CAR does not. The formula for calculating the BHAR is usually given by

$$B\hat{H}AR_i = [\Pi_{t=T_1}^{T_2}(1 + R_{it}) - 1] - [\Pi_{t=T_1}^{T_2}(1 + E(R_{it})) - 1] \tag{14.21}$$

where $E(R_{it})$ is the expected return. Usually, when constructing BHARs the expected return is based on a non-event firm or portfolio of firms that is matched in some way to the event firm (e.g. based on size, industry, etc.). Alternatively, although less desirably, it could be obtained from a benchmark such as a stock market index.

If desired, we can then sum the $BHAR_i$ across the N firms to construct an aggregate measure. BHARs have been advocated, among others, by Barber and Lyon (1997) and Lyon *et al.* (1999) because they better match the 'investor experience' than CARs given the former's use of geometric rather than arithmetic averaging. CARs represent biased estimates of the actual returns received by investors. However, by contrast, Fama (1998) in particular argues in favour of the use of CARs rather than BHARs. The latter seem to be more adversely affected by skewness in the sample of abnormal returns than the former because of the impact of compounding in BHARs.[5] In addition, Fama indicates that the average CAR

[5] Although Lyon *et al.* (1999) propose a skewness-adjusted t-statistic with bootstrapping to mitigate this problem.

increases at a rate of $(T_2 - T_1)$ with the number of months included in the sum, whereas its standard error increases only at a rate $\sqrt{(T_2 - T_1)}$. This is not true for BHARs where the standard errors grow at the faster rate $(T_2 - T_1)$ rather than its square root. Hence any inaccuracies in measuring expected returns will be more serious for BHARs as another consequence of compounding.

Event time versus calendar time analysis

All of the procedures discussed above have involved conducting analysis in *event time*. There is, however, an alternative approach that involves using *calendar time*, advocated by Fama (1998) and Mitchell and Stafford (2000) among others. In essence, using a calendar time methodology involves running a time series regression and examining the intercept from that regression. The dependent variable is a series of portfolio returns, which measure the average returns at each point in time of the set of firms that have undergone the event of interest within a pre-defined measurement period before that time. So, for example, we might choose to examine the returns of firms for a year after the event that they announce cessation of their dividend payments. Then, for each observation t, the dependent variable will be the average return on all firms that stopped paying dividends at any point during the past year. One year after the event, by construction the firm will drop out of the portfolio. Hence the number of firms within the portfolio will vary over time (as the number of firms ceasing dividend payment varies) and the portfolio will effectively be rebalanced each month. The explanatory variables may be risk measures from the Carhart (1997) four-factor model for example – this will be discussed in detail below.

The calendar time approach will weight each time period equally and thus the weight on each individual firm in the sample will vary inversely with the number of other firms that have undergone the event during the observation period. This may be problematic and will result in a loss of power to detect an effect if managers time their events to take advantage of misvaluations.

Small samples and non-normality

The test statistics presented in the previous section are all asymptotic, and problems may arise either if the estimation window (T) is too short, or if the number of firms (N) is too small when the firm-aggregated statistic is used. As we discussed earlier in the book, it is well known that stock returns are leptokurtic and tend to have longer lower tails than upper tails. And particularly with small samples, the presence of outliers – for example, very large returns during the estimation window affecting the market model parameter estimation or the residual variance estimates – may also be problematic. One possible remedy would be to use a bootstrap approach to computing the test statistics.

A second strategy for dealing with non-normality would be to use a non-parametric test. Such tests are robust in the presence of non-normal distributions, although they are usually less powerful than their parametric counterparts. In the present context, we could test the null hypothesis that the proportion of positive abnormal returns is not affected by the event. In other words, the proportion of positive abnormal returns across firms remains at the expected level. We could then use the test statistic, Z_p

$$Z_p = \frac{[p - p^*]}{[p^*(1 - p^*)/N]^{1/2}} \tag{14.22}$$

where p is the actual proportion of negative abnormal returns during the event window and p^* is the expected proportion of negative abnormal returns. Under the null hypothesis, the test statistic follows a binomial distribution, which can be approximated by the standard normal distribution. Sometimes p^* is set to 0.5, but this may not be appropriate if the return distribution is skewed, which is typically the case. Instead, it is better to calculate p^* based on the proportion of negative abnormal returns during the estimation window. The Wilcoxon signed-rank test can also be used.

Event studies – some further issues

A further implicit assumption in the standard event test methodology is that the events themselves occur involuntarily. In practice, however, firms often have discretion about the extent, timing and presentational forms of the announcements that they make. Thus they are likely to use any discretion they have to make announcements when market reactions are going to be the most favourable. For example, where the local regulatory rules allow discretion, firms may release bad news when the markets are closed or when the media and investors are preoccupied with other significant news items. Prabhala (1997) discusses the implications of and solutions to the endogeneity of the firm's decision about when (and perhaps even whether) to make an announcement. When a firm chooses not to announce at a particular time, we have a sort of truncated sample since we can observe events only for firms who choose to make an announcement.

A way of simultaneously dealing with a number of the issues highlighted above (i.e. differing return variances across firms, changing return variances over time, and clustering of events across firms) is to use what has been termed generalised least squares (GLS) in constructing the test statistics. In essence this works by constructing a variance-covariance matrix from the abnormal returns and using this to weight the returns in computing the aggregate test statistic – see Armitage (1995) for further details.

We can see from the above that a range of procedures exists for conducting event studies. The core of the approach is the same in each case, but they differ according to how the aggregation is performed over time and across firms and this affects the method of calculation of the standard deviations. So how do we choose which approach to use? Hopefully, given the context and the nature of the events under consideration, we can gain a reasonable idea of which approach is likely to be the most justifiable. For example, is clustering an issue? Is it expectable that the return variances will have changed over time? Is it important to allow for the variances of returns to vary between firms? By answering these questions, we can usually select the appropriate procedure. But if in doubt, it is always advisable to examine a range of methods and to compare the results as a robustness check. With luck, the various calculation techniques will lead to the same conclusion.

14.9.4 Conducting an event study using Excel

This section will now use the core of the approaches described above in order to conduct an event study. While this ought to be enough to get started and to obtain some indicative results, it is important to note that there is far more that can be done with event studies to make them more rigorous than the approach presented here and readers are encouraged to consult the papers cited above for further details.

The first step would be to decide which event to consider the impact of, and there is certainly no shortage of possibilities (dividend announcements; stock spit announcements; index composition changes; merger announcements; CEO turnover; new contract announcements; macroeconomic announcements, etc.). Once this is done and the data are collected, the time-consuming part is to then organise them in a way to make them easy to work with. It would be possible to conduct the analysis in any software package for data analysis, including EViews. However, since the bulk of the task involves data arrangement and the econometric part is usually not sophisticated (in most cases, a regression will not even be conducted), it probably makes sense to revert back to Microsoft Excel or a similar spreadsheet package.[6]

The starting point for the analysis conducted here are the abnormal returns for $N = 20$ firms, which are given in the Excel file 'Event.xls', and have already been calculated using the market model using equations (14.1) and (14.2). The returns are given for days -259 to $+263$. The raw data are on the sheet 'abnormal returns'. The spreadsheet has been set up with the data are aligned on the event day, so while the firms underwent the event on different days, the spreadsheet is constructed so that day '0' is the event day in the same row for all firms. The estimation period is from day –259 to day –10 inclusive (249 days), while the event periods examined are $(T-10, T-1)$, day T itself, $(T+1, T+10)$ and $(T+1, T+250)$. The first of these windows allows us to examine whether there was any leakage of information that affected stock returns prior to the event. Whether there is an immediate effect on the day that the event occurs will depend on whether the announcement is made in advance or it is a 'surprise' to the markets. If the event was known in advance to be happening on day T then the impact on the markets that day may be muted since it could have already been reflected in prices. Note that in this case the adjustment in equation (14.4) is not employed since the estimation period ($T = 249$) is quite long and would render the adjustment term negligible.

We first calculate the average return across all twenty firms for each day during the estimation and event windows in column V of the 'abnormal returns' sheet using the Excel AVERAGE formula in the usual way. All of the calculations of the

[6] The example below uses a small sample of real data from a real event, but no details are given as to the nature of the event so that they can be distributed freely with the book.

key statistics are done on a separate sheet which I have called 'summary stats'. The sheet first calculates the AR for day T and the CARs for the date ranges using equations (14.1) and (14.6) respectively for each individual firm and also for the average across all firms.

The next step is to calculate the variances of the abnormal returns or cumulative abnormal returns. For day T, this is done using equation (14.3), which is simply the time series variance of returns during the estimation window and placed in row 2 (and copied directly into row 11). For the multi-day event windows, the one-day variance from (14.3) is scaled up by the number of days in the event window (10 or 250) using equation (14.7). Then the test statistics are calculated by dividing the AR by its respective standard deviation (i.e. the square root of the variance) using (14.5) or its CAR equivalent in (14.8). Finally, the easiest way to obtain p-values for the tests is to use the TDIST function in Excel for a two-sided test and with a large number of degrees of freedom (say, 1,000), so that it approximates the normal distribution.

As discussed in the previous section, there are several potential issues with the fairly simple event study methodology just described. So, for robustness, it is a good idea to examine different ways of tackling the problem, and two possible checks are given in columns X and Y of the 'summary stats' sheet. Both of these procedures can only be undertaken based on the average across firms and not at the individual firm level. The first tweak is to calculate the standard deviation used in the test statistics cross-sectionally in order to allow for the possibility that the return variances may have changed (typically, risen) around the time of the event. Thus we simply take the variance across firms for the abnormal return or cumulative abnormal return of interest, divide this by N (i.e. 20) and then proceed in the usual way.

A further possibility examined in column Y is to equally weight firms by calculating the average of the standardised abnormal returns as in equation (14.19) or (14.20). Then the test statistic is simply this average multiplied by the square root of the N.

If we now consider the results on this sheet, it is clear that there is little evidence of a short-term reaction to the event. During the two trading weeks before the event, ($T - 10$ to $T - 1$), only one firm has a significant abnormal return at the 5% level (firm 20 has a CAR of 15.43% with a test statistic of 2.02). None of the individual firms have significant returns on the event date (T), and neither do any of them show significance in the short post-event window ($T + 1$ to $T + 10$). It is over the longer term – the next trading year – where there is some action. Now five firms have statistically significant returns together with economically quite large cumulative abnormal returns of 20% to 55%.

Examining the aggregate-level results, it is reassuring that the three slightly different approaches in columns W to Y yield very similar conclusions. Here the null hypothesis is that the average abnormal return (or average cumulative abnormal return) is zero. There is again no discernible market reaction before, on, or in the short-run after, the event. However, the long-run abnormal return is positive and highly statistically significant whichever of the three approaches is considered. Interestingly, the variance estimates before the event (at times $t - 10$

to $T-1$) are higher for the cross-sectional approach in (14.18), although they are lower for cross-sectional approach during and after the event.

Finally, in the third sheet of the Event.xls workbook, labelled 'non-parametric test', the non-parametric statistic Z_p of equation (14.22) is calculated and then the p-value is obtained using the TDIST function as above. This examines the null hypothesis that the proportion of abnormal returns around the event is the same as it was during the estimation window. So the first calculation row of the sheet (row 2) calculates p^*, the expected proportion of negative returns based on data from the estimation window. Then for each event period range, we calculate p, the actual proportion of negative returns.[7]

The expected proportion of negative returns varies from 0.43 for firm 18 to 0.55 for firm 8, but the actual proportions for the short pre- and post-event windows are often much lower than that. For example, for firm 1, p was 0.3 (i.e. negative returns on only three days from ten) before the event. Pre-event, six of the twenty firms have significant differences between p and p^*, while for the two weeks immediately after the event, only three of them show significant differences. Over the long-run, however, there are no significant differences between the expected and actual proportions of negative return days – either for any of the individual firms or for the average.

14.10 Tests of the CAPM and the Fama–French Methodology

14.10.1 Testing the CAPM

The basics

Before moving on to the more sophisticated multi-factor models, it may be useful to review the standard approach that was developed for testing the CAPM. This is not the place for a detailed discussion of the motivation for the CAPM or its derivation – such a discussion can be found at an accessible level in Bodie *et al.* (2011) or most other finance textbooks; alternatively, see Campbell *et al.* (1997) for a more technical treatment. A good introduction to the general area of asset pricing tests is given in the book by Cuthbertson and Nitzsche (2004).

The most commonly quoted equation for the CAPM is

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \tag{14.23}$$

So the CAPM states that the expected return on any stock i is equal to the risk-free rate of interest, R_f, plus a risk premium. This risk premium is equal to the risk premium per unit of risk, also known as the market risk premium, $[E(R_m) - R_f]$, multiplied by the measure of how risky the stock is, known as 'beta', β_i. Beta is not observable from the market and must be calculated, and hence tests of the CAPM are usually done in two steps – first, estimating the stock betas and second, actually testing the model. It is important to note that the CAPM is an equilibrium model,

[7] Note of course that it is not possible to calculate Z for the event date by itself since the proportion of negative returns, p would be either exactly zero or exactly one.

or a model in terms of expectations. Thus, we would not expect the CAPM to hold in every time period for every stock. But if it is a good model, then it should hold 'on average'. Usually, we will use a broad stock market index as a proxy for the market portfolio and the yield on short-term Treasury bills as the risk-free rate.

A stock's beta can be calculated in two ways – one approach is to calculate it directly as the covariance between the stock's *excess* return and the excess return on the market portfolio, divided by the variance of the excess returns on the market portfolio

$$\beta_i = \frac{Cov(R_i^e, R_m^e)}{Var(R_m^e)} \tag{14.24}$$

where the ^{e} superscript denotes excess returns (i.e. the return with the risk-free rate subtracted from it). Alternatively, and equivalently, we can run a simple *time series* regression of the excess stock returns on the excess returns to the market portfolio separately for each stock, and the slope estimate will be the beta

$$R_{i,t}^e = \alpha_i + \beta_i R_{m,t}^e + u_{i,t}, \qquad i = 1, \ldots, N; \quad t = 1, \ldots, T \tag{14.25}$$

where N is the total number of stocks in the sample and T is the number of time series observations on each stock.

The intercept estimate ($\hat{\alpha}_i$) from this regression would be 'Jensen's alpha' for the stock, which would measure how much the stock underperformed or outperformed what would have been expected given its level of market risk. It is probably not very interesting to examine the alpha for an individual stock, but we could use exactly the same regression to test the performance of portfolios, trading strategies and so on – all we would do would be to replace the excess returns that comprise the dependent variable with those from the portfolio or trading rule.

Returning to testing the CAPM, suppose that we had a sample of 100 stocks ($N = 100$) and their returns using five years of monthly data ($T = 60$). The first step would be to run 100 time series regressions (one for each individual stock), the regressions being run with the sixty monthly data points. Then the second stage would involve a single cross-sectional regression of the average (over time) of the stock returns on a constant and the betas

$$\bar{R}_i = \lambda_0 + \lambda_1 \beta_i + v_i, \qquad i = 1, \ldots, N \tag{14.26}$$

where $\bar{R}_i$ is the return for stock i averaged over the sixty months. Notice that, unlike the first stage, this second stage regression now involves the actual returns and not excess returns. Essentially, the CAPM says that stocks with higher betas are more risky and therefore should command higher average returns to compensate investors for that risk.

If the CAPM is a valid model, two key predictions arise which can be tested using this second stage regression: $\lambda_0 = R_f$ and $\lambda_1 = [R_m - R_f]$. So, to find support for the CAPM, we would expect to see the intercept estimate being close to the risk-free rate of interest and the slope being close to the market risk premium.

Two further implications of the CAPM being valid are first, that there is a linear relationship between a stock's return and its beta and second, that no other variables should help to explain the cross-sectional variation in returns. So, in other words, any additional variable we add to the second stage regression (14.26) should not have a statistically significant parameter estimate attached to it. We could thus for example run the augmented regression

$$\bar{R}_i = \lambda_0 + \lambda_1 \beta_i + \lambda_2 \beta_i^2 + \lambda_3 \sigma_i^2 + v_i \tag{14.27}$$

where β_i^2 is the squared beta for stock i and σ_i^2 is the variance of the residuals from the first stage regression, which is a measure of idiosyncratic risk for stock i. The squared beta term can capture whether there is any non-linearity in the relationship between returns and beta. If the CAPM is a valid and complete model, then we should see that $\lambda_2 = 0$ and $\lambda_3 = 0$.

However, research has indicated that the CAPM is not a complete model of stock returns. In particular, it has been found that returns are systematically higher for small capitalisation stocks than the CAPM would predict, and similarly, returns are systematically higher for 'value' stocks (those with low market-to-book or price-to-earnings ratios) than the CAPM would predict. We can test this directly using a different augmented second stage regression

$$\bar{R}_i = \lambda_0 + \lambda_1 \beta_i + \lambda_2 MV_i + \lambda_3 BTM_i + v_i \tag{14.28}$$

where MV_i is the market capitalisation for stock i and BTM_i is is the ratio of its book value to its market value of equity.[8] This is the kind of model employed by Fama and French (1992), as discussed below. As for equation (14.27), the test for the CAPM to be supported by the data would be $\lambda_2 = 0$ and $\lambda_3 = 0$.

Unfortunately, returns data are beset by problems that can render the results from tests of the CAPM dubious or possibly even invalid. First, the familiar non-normality of returns can lead to problems with tests in finite samples – while normality is not a specific theoretical requirement of the CAPM, it is required for valid hypothesis testing. Second, there is also likely to be heteroscedasticity in the returns. More recent research testing the CAPM has used the generalised method of moments (GMM), where estimators can be constructed that are robust to these issues – see for, example, Cochrane (2005). A final important problem is the measurement error in beta discussed extensively in section 5.13 of this book. In order to minimise such measurement errors, the beta estimates can be based on portfolios rather than individual securities. Alternatively, the Shanken (1992) correction can be applied, where the standard deviation in the test statistic is multiplied by a factor to adjust for the measurement error.

The Fama–MacBeth approach

Fama and MacBeth (1973) used the two stage approach to testing the CAPM outlined above, but using a *time series of cross-sections*. The basics are exactly as described above, but instead of running a single time-series regression for each

[8] Note that many studies use the market-to-book ratio, which is simply one divided by the book-to-market ratio – so value stocks have a low number for the former and a high number for the latter.

stock and then a single cross-sectional regression, the estimation is conducted with a rolling window.

Fama and MacBeth employ five years of observations to estimate the CAPM betas and the other risk measures (i.e. the standard deviation and squared beta) and these are used as the explanatory variables in a set of cross-sectional regressions each month for the following four years. The estimation period is then rolled forward four years and the process continues until the end of the sample period is reached.[9] To illustrate, their initial time series estimation period for the betas is January 1930 to December 1934. The cross-sectional regressions are run with monthly returns on each stock as the dependent variable for January 1935, and then separately for February 1935, . . . , to December 1938. The sample is then rolled forward with the beta estimation from January 1934 to December 1938 and the cross-sectional regressions now beginning January 1939. In this way, they end up with a cross-sectional regression for every month in the sample (except for the first five years used for the initial beta estimations).

Since we will have one estimate of the lambdas, $\hat{\lambda}_{j,t}$ $(j = 1, 2, 3, 4)$, for each time period t, we can form a t-ratio for each of these as being the average over t, denoted $\hat{\lambda}_j$, divided by its standard error (which is the standard deviation over time divided by the square root of the number of time series estimates of the $\hat{\lambda}_{j,t}$).

Thus the average value over t of $\hat{\lambda}_{j,t}$ can be calculated as

$$\hat{\lambda}_j = \frac{1}{T_{FMB}} \sum_{t=1}^{T_{FMB}} \hat{\lambda}_{j,t}, \quad j = 1, 2, 3, 4 \tag{14.29}$$

where T_{FMB} is the number of cross-sectional regressions used in the second stage of the test, and the standard deviation is

$$\hat{\sigma}_j = \sqrt{\frac{1}{T_{FMB} - 1} \sum_{t=1}^{T_{FMB}} (\hat{\lambda}_{j,t} - \hat{\lambda}_j)^2} \tag{14.30}$$

The test statistic is then simply $\sqrt{T_{FMB}}\hat{\lambda}_j / \hat{\sigma}_j$, which is asymptotically standard normal, or follows a t distribution with $T_{FMB} - 1$ degree of freedom in finite samples. The key results from Fama and MacBeth corroborate other early evidence by Black, Jensen and Scholes (1972), and are summarised in table 14.3.

We can compare the estimated values of the intercept and slope with the actual values of the risk-free rate (R_f) and the market risk premium $[\bar{R}_m - \bar{R}_f]$, which are, for the full-sample corresponding to the results presented in the table, 0.013 and 0.143 respectively. The parameter estimates $\hat{\lambda}_0$ and $\hat{\lambda}_1$ have the correct signs (both are positive). Thus the implied risk-free rate is positive and so is the relationship between returns and beta – both parameters are significantly different from zero, although they become insignificant when the other risk measures are

[9] The main reason that the updating was only undertaken every four years was due to the lack of computing power available at that time. More recent studies would do this annually or even monthly.

Table 14.3 Fama and MacBeth's results on testing the CAPM

Model	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$
Model 1: CAPM	0.0061*	0.0085*		
	(3.24)	(2.57)		
Model 2: Augmented CAPM	0.0020	0.0114	−0.0026	0.0516
	(0.55)	(1.85)	(−0.86)	(1.11)

Notes: *t*-ratios in parentheses; * denotes significance at the 5% level.
Source: Fama and MacBeth (1973), numbers extracted from their Table 3.

included as in the second row of the table. Hence it has been argued that there is *qualitative* support for the CAPM but not *quantitative* support as the intercept and slope are not of the appropriate sizes, although the differences between the estimated parameters and their expected values are not statistically significant for Fama and MacBeth's whole sample. It is also worth noting from the second row of the table that squared beta and idiosyncratic risk have parameters that are even less significant than beta itself in explaining the cross-sectional variation in returns.

14.10.2 Asset pricing tests – the Fama–French approach

Of all of the approaches to asset pricing tests that have been developed, the range of techniques pioneered by Fama and French in a series of papers is by far the most commonly employed. The 'Fama–French methodology' is not really a single technique but rather a family of related approaches based on the notion that market risk is insufficient to explain the cross-section of stock returns – in other words, why some stocks generate higher average returns than others.

The Fama–French and Carhart models, described in detail below, seek to measure abnormal returns after allowing for the impact of the characteristics of the firms or portfolios under consideration. It is well-established in the finance literature that certain types of stocks yield, on average, considerably higher returns than others. For example, the stocks of small companies, value stocks (those with low price-to-earnings ratios), and stocks with momentum (that have experienced recent price increases), typically yield higher returns than those having the opposite characteristics. This has important implications for asset pricing and for the way that we think about risk and expected returns. If, for example, we wanted to evaluate the performance of a fund manager, it would be important to take the characteristics of these portfolios into account to avoid incorrectly labelling a manager as having stock-picking skills when he routinely followed a strategy of buying small, value stocks with momentum, which will on average outperform the equity market as a whole.

Fama–French (1992)

The Fama–French (1992) approach, like Fama and MacBeth (1973), is based on a time series of cross-sections model. Here, we run a set of cross-sectional regressions of the form

$$R_{i,t} = \alpha_{0,t} + \alpha_{1,t}\beta_{i,t} + \alpha_{2,t}MV_{i,t} + \alpha_{3,t}BTM_{i,t} + u_{i,t} \tag{14.31}$$

where $R_{i,t}$ are again the monthly returns, $\beta_{i,t}$ are the CAPM betas, $MV_{i,t}$ are the market capitalisations, and $BTM_{i,t}$ are the book-to-price ratios, each for firm i and month t. So the explanatory variables in the regressions here are the firm characteristics themselves. Fama and French show that when we employ size and book-to-market in the cross-sectional regressions, these are highly significantly related to returns (with negative and positive signs respectively) so that small and value stocks earn higher returns all else equal than growth or large stocks. They also show that market beta is not significant in the regression (and even has the wrong sign), providing very strong evidence against the CAPM.

Fama–French (1993)

Fama and French (1993) use a factor-based model in the context of a time series regression which is now run separately on each portfolio i

$$R_{i,t} = \alpha_i + \beta_{i,M}RMRF_t + \beta_{i,S}SMB_t + \beta_{i,V}HML_t + \epsilon_{i,t} \tag{14.32}$$

where $R_{i,t}$ is the return on stock or portfolio i at time t, $RMRF$, SMB, and HML are the *factor mimicking portfolio* returns for market excess returns, firm size, and value respectively.[10]

The factor mimicking portfolios are designed to have unit exposure to the factor concerned and zero exposure to all other factors. In more detail, the factors in the Fama and French (1993) model are constructed as follows. The excess market return is measured as the difference in returns between the S&P500 index and the yield on Treasury bills ($RMRF$); SMB is the difference in returns between a portfolio of small stocks and a portfolio of large stocks, termed 'small minus big' portfolio returns; HML is the difference in returns between a portfolio of value stocks with high book-value to market-value ratios and a portfolio of growth stocks with low book-value to market-value ratios, termed 'high minus low' portfolio returns. One of the main reasons they use factor-mimicking portfolios rather than continuing their (1992) approach is that they want to also include bonds in the set of asset returns considered, and these do not have obvious analogues to market capitalisation or the book-to-market ratio.

In Fama and French's (1993) case, these time series regressions are run on portfolios of stocks that have been two-way sorted according to their book-to-market ratios and their market capitalisations. It is then possible to compare the parameter estimates qualitatively across the portfolios i. The parameter estimates from these time series regressions are known as *factor loadings* that measure the

[10] While this model could be applied to individual stocks, it makes more sense in the context of portfolios, although the principles are the same.

sensitivity of each individual portfolio to each of the factors. We will obtain a separate set of factor loadings for each portfolio i since each portfolio is the subject of a different time series regression and will have different sensitivities to the risk factors. Fama and French (1993) qualitatively compare these factor loadings across a set of twenty-five portfolios that have been two-way sorted on their size and book-to-market ratios.

Then, the second stage in this approach is to use the factor loadings from the first stage as explanatory variables in a cross-sectional regression

$$\bar{R}_i = \alpha + \lambda_M \beta_{i,M} + \lambda_S \beta_{i,S} + \lambda_V \beta_{i,V} + e_i \tag{14.33}$$

We can interpret the second stage regression parameters, λ_M, λ_S, λ_V as *factor risk premia* – in other words, they show the amount of extra return that is generated on average from taking on an additional unit of that source of risk.

Since the factor loadings and risk premia have a tendency to vary over time, the model is estimated using a rolling window. For example, the time series model in equation (14.32) is typically estimated using five years of monthly data, and then the λs would be estimated from equation (14.33) using separate cross-sectional regressions with a monthly return for each of the following twelve months. The sample would then be rolled forward by a year with a new set of βs being estimated from (14.32) and then a new set of twelve estimates of λ produced and so on. Alternatively, the rolling update could occur monthly. Either way, there will be one estimate of each of the λs for every month after the initial five-year beta estimation window, which we would then average to get the overall estimates of the risk premia.

Fama and French (1993) apply the model to their twenty-five size- and value-sorted portfolios and argue that the statistical significance of the lambdas in the second stage regressions and the high R^2 values are indicative of the importance of size and value as explanators of the cross-sectional variation in returns.

Carhart (1997)

Since Carhart's (1997) study on mutual fund performance persistence, it has become customary to add a fourth factor to the equations above based on momentum, measured as the difference between the returns on the best performing stocks over the past year and the worst performing stocks – this factor is known as UMD – 'up-minus-down'. Equation (14.32) then becomes

$$R_{i,t} = \alpha_i + \beta_{i,M} RMRF_t + \beta_{i,S} SMB_t + \beta_{i,V} HML_t + \beta_{i,U} UMD_t + \epsilon_{i,t} \tag{14.34}$$

And, if desired, equation (14.33) becomes[11].

$$\bar{R}_i = \alpha + \lambda_M \beta_{i,M} + \lambda_S \beta_{i,S} + \lambda_V \beta_{i,V} + \lambda_U \beta_{i,U} + e_i \tag{14.35}$$

[11] Note that Carhart's paper does not use this second-stage cross-sectional regression containing the factor sensitivities.

Carhart forms decile portfolios of mutual funds based on their one-year lagged performance and runs the time series regression of equation (14.34) on each of them. He finds that the mutual funds which performed best last year (in the top decile) also had positive exposure to the momentum factor (UMD) while those which performed worst had negative exposure. Hence a significant portion of the momentum that exists at the fund level arises from momentum in the stocks that those funds are holding.

14.10.3 The Fama–MacBeth procedure in EViews

It should be clear from the discussion above that there is nothing particularly complex about the two-stage procedure – it only involves two sets of standard linear regressions. The hard part is really in collecting and organising the data. If we wished to do a more sophisticated study – for example, using a bootstrapping procedure or using the Shanken correction, this would require more analysis than is conducted in the illustration below. However, hopefully the EViews code and explanation will be sufficient to demonstrate how to apply the procedures to any set of data.

The example employed here is taken from the study by Gregory, Tharyan and Chistidis (2013) that examines the performance of several different variants of the Fama–French and Carhart models using the Fama–MacBeth methodology in the UK following several earlier studies showing that these approaches appear to work far less well for the UK than the US. The data required are provided by Gregory *et al.* on their web site.[12] Note that their data have been refined and further cleaned since their paper was written (i.e. the web site data are not identical to those used in the paper) and as a result the parameter estimates presented here deviate slightly from theirs. However, given that the motivation for this exercise is to demonstrate how the Fama–MacBeth approach can be used in EViews, this difference should not be consequential. The two data files used are 'monthlyfactors.csv' and 'vw_sizebm_25groups.csv'. The former file includes the time series of returns on all of the factors (SMB, HML, UMD, RMRF, the return on the market portfolio (RM) and the return on the risk-free asset (RF)), while the latter includes the time series of returns on twenty-five value-weighted portfolios formed from a large universe of stocks, two-way sorted according to their sizes and book-to-market ratios.

The first step in this analysis for conducting the Fama–French or Carhart procedures using the methodology developed by Fama and MacBeth is to create a new EViews workfile which I have called 'ff_example.wf1' and to import the two csv data files into it. The data in both cases run from October 1980 to December 2012, making a total of 387 data points. However, in order to obtain results as close as possible to those of the original paper, when running the regressions, the period is from October 1980 to December 2010 (363 data points). We then need to set up a program file along the lines of those set up in the previous chapter – I have called mine 'FF_PROG.prg'.

[12] http://business-school.exeter.ac.uk/research/areas/centres/xfi/research/famafrench/files.

The full code to run the tests is as follows, and annotated below.

```
'READ DATA
LOAD C:\ CHRIS\ BOOK\ BOOK3E\ DATA\ FF_EXAMPLE.WF1

'TRANSFORM ACTUAL RETURNS INTO EXCESS RETURNS
SL=SL-RF
S2=S2-RF
S3=S3-RF
S4=S4-RF
SH=SH-RF
S2L=S2L-RF
S22=S22-RF
S23=S23-RF
S24=S24-RF
S2H=S2H-RF
M3L=M3L-RF
M32=M32-RF
M33=M33-RF
M34=M34-RF
M3H=M3H-RF
B4L=B4L-RF
B42=B42-RF
B43=B43-RF
B44=B44-RF
B4H=B4H-RF
BL=BL-RF
B2=B2-RF
B3=B3-RF
B4=B4-RF
BH=BH-RF

'DEFINE THE NUMBER OF TIME SERIES OBSERVATIONS
!NOBS=363

'CREATE SERIES TO PUT BETAS FROM STAGE 1
'AND LAMBDAS FROM STAGE 2 INTO
SERIES BETA_C
SERIES BETA_RMRF
SERIES BETA_UMD
SERIES BETA_HML
SERIES BETA_SMB
SERIES LAMBDA_C
SERIES LAMBDA_RMRF
SERIES LAMBDA_UMD
SERIES LAMBDA_HML
```

```
SERIES LAMBDA_SMB
SERIES LAMBDA_R2
SCALAR LAMBDA_C_MEAN
SCALAR LAMBDA_C_TRATIO
SCALAR LAMBDA_RMRF_MEAN
SCALAR LAMBDA_RMRF_TRATIO
SCALAR LAMBDA_UMD_MEAN
SCALAR LAMBDA_UMD_TRATIO
SCALAR LAMBDA_HML_MEAN
SCALAR LAMBDA_HML_TRATIO
SCALAR LAMBDA_SMB_MEAN
SCALAR LAMBDA_SMB_TRATIO
SCALAR LAMBDA_R2_MEAN
'THIS LOOP CREATES THE SERIES TO PUT THE
'CROSS-SECTIONAL DATA IN
FOR !M = 1 TO 387
SERIES TIME{%M}
NEXT

'NOW RUN THE FIRST STAGE TIME-SERIES REGRESSIONS
'SEPARATELY FOR EACH PORTFOLIO AND
'PUT THE BETAS INTO THE APPROPRIATE SERIES
SMPL 1980:10 2010:12
!J=1
FOR %Y SL S2 S3 S4 SH S2L S22 S23 S24 S2H M3L M32 M33 M34 M3H
   B4L B42 B43 B44 B4H BL B2 B3 B4 BH
'THE PREVIOUS COMMAND WITH VARIABLE NAMES
'NEEDS TO ALL GO ON ONE LINE
EQUATION EQ1.LS {%Y} C RMRF UMD HML SMB
BETA_C(!J)=@COEFS(1)
BETA_RMRF(!J)=@COEFS(2)
BETA_UMD(!J)=@COEFS(3)
BETA_HML(!J)=@COEFS(4)
BETA_SMB(!J)=@COEFS(5)
!J=!J+1
NEXT

'NOW RESORT THE DATA SO THAT EACH COLUMN IS A
'MONTH AND EACH ROW IS RETURNS ON PORTFOLIOS
FOR !K=1 TO 387
TIME!K(1)=SL(!K)
TIME!K(2)=S2(!K)
TIME!K(3)=S3(!K)
TIME!K(4)=S4(!K)
TIME!K(5)=SH(!K)
TIME!K(6)=S2L(!K)
```

```
TIME!K(7)=S22(!K)
TIME!K(8)=S23(!K)
TIME!K(9)=S24(!K)
TIME!K(10)=S2H(!K)
TIME!K(11)=M3L(!K)
TIME!K(12)=M32(!K)
TIME!K(13)=M33(!K)
TIME!K(14)=M34(!K)
TIME!K(15)=M3H(!K)
TIME!K(16)=B4L(!K)
TIME!K(17)=B42(!K)
TIME!K(18)=B43(!K)
TIME!K(19)=B44(!K)
TIME!K(20)=B4H(!K)
TIME!K(21)=BL(!K)
TIME!K(22)=B2(!K)
TIME!K(23)=B3(!K)
TIME!K(24)=B4(!K)
TIME!K(25)=BH(!K)
NEXT

'RUN 2ND STAGE CROSS-SECTIONAL REGRESSIONS
FOR !Z = 1 TO !NOBS
EQUATION EQ1.LS TIME!Z C BETA_RMRF BETA_UMD BETA_HML
   BETA_SMB
LAMBDA_C(!Z)=@COEFS(1)
LAMBDA_RMRF(!Z)=@COEFS(2)
LAMBDA_UMD(!Z)=@COEFS(3)
LAMBDA_HML(!Z)=@COEFS(4)
LAMBDA_SMB(!Z)=@COEFS(5)
LAMBDA_R2(!Z)=@R2
NEXT

'FINALLY, ESTIMATE THE MEANS AND T-RATIOS
'FOR THE LAMBDA ESTIMATES IN THE SECOND STAGE
LAMBDA_C_MEAN =@MEAN(LAMBDA_C)
LAMBDA_C_TRATIO=@SQRT(!NOBS)*@MEAN(LAMBDA_C)/
   @STDEV(LAMBDA_C)
LAMBDA_RMRF_MEAN=@MEAN(LAMBDA_RMRF)
LAMBDA_RMRF_TRATIO=@SQRT(!NOBS)*@MEAN(LAMBDA_
   RMRF)/@STDEV(LAMBDA_RMRF)
LAMBDA_UMD_MEAN=@MEAN(LAMBDA_UMD)
LAMBDA_UMD_TRATIO=@SQRT(!NOBS)*@MEAN(LAMBDA_
   UMD)/@STDEV(LAMBDA_UMD)
LAMBDA_HML_MEAN=@MEAN(LAMBDA_HML)
```

```
LAMBDA_HML_TRATIO=@SQRT(!NOBS)*@MEAN(LAMBDA_
   HML)/@STDEV(LAMBDA_HML)
LAMBDA_SMB_MEAN=@MEAN(LAMBDA_SMB)
LAMBDA_SMB_TRATIO=@SQRT(!NOBS)*@MEAN(LAMBDA_SMB)/
   @STDEV(LAMBDA_SMB)
LAMBDA_R2_MEAN=@MEAN(LAMBDA_R2)
```

We can think of this program as comprising of several sections. The first step is to transform all of the raw portfolio returns into excess returns which are required to compute the betas in the first stage of Fama–MacBeth. This is fairly simple to do and writes over the original series with their excess return counterparts.

The line (!NOBS=363) ensures that the same sample period as the paper by Gregory *et al.* is employed throughout. The next stage involves creating the arrays to put the betas and lambdas in. These are set up as series since there will be one entry for each regression. Then we also need the final estimates for each of the lambda parameters, which will be the time series averages of the cross-sections.

We need to first run a set of time series regressions to estimate the betas but we will later need to estimate a set of cross-sectional regressions. This presents a problem because the data can only be organised in one way or the other in EViews. So the following three lines

```
FOR !M = 1 TO 387
SERIES TIME{M}
NEXT
```

set up a set of 387 new series called TIME1, TIME2, . . . , TIME387 which we will subsequently organise as cross-sectional data. !M in curly brackets is what tells EViews to add the numbers 1, 2, . . . , onto the word TIME to create the names for the new series. These three lines of code very efficiently replace 387 separate lines of code that we would otherwise have had to have written such as SERIES TIME1 etc.

Then we set up and run all of the first-stage time series regressions. We want to run the Carhart 4-factor model separately for each of the twenty-five portfolios. It would be possible to have twenty-five separate regression statements, but it seems easier and more efficient to set these up in a loop. SMPL 1980:10 2010:12 runs the regressions for the period 1980:10 to 2010:12 only rather than on the whole sample period.

```
The statements
FOR %Y followed by the list of variable names
. . .
NEXT
```

constitute the main loop that runs over all the twenty-five series. Then the line

```
EQUATION EQ1.LS {%Y} C RMRF UMD HML SMB
```

runs an OLS time series regression for each of the twenty-five series in the loop defined in the previous line on a constant and the four variables RMRF UMD HML SMB. This effectively uses equation (14.34) from above. We need to store the estimates from these regressions into separate series for each parameter. The lines beginning BETA_C(!J)=@COEFS(1) do this. The letter J is an index which is defined outside the loop to start with a value of 1 (the statement !J=1) and then as each regression is run, the value of J is increased by 1 (the statement !J=!J+1 does this). So, the loop starts off with J=1 and the regression will be run with the series SL as the dependent variable. The the intercept (i.e. the alpha) from this regression will be placed as the first entry in BETA_C (i.e. it will be BETA_C(1)), the parameter estimate on the RMRF term will be placed in BETA_RMRF(1) and so on. Then the value of J will be increased by 1 to 2, and the second regression will have dependent variable S2. Its intercept estimate will be placed in BETA_C(2), the slope estimate on RMRF will be placed in BETA_RMRF(2) and so on. This will continue until the final regression is run on the twenty-fifth series, which will be BH, with its intercept estimate being placed in BETA_C(25). We should thus note that while these BETA_ series were set up with the total number of observations in the workfile (i.e. they will have 387 rows), only the first twenty-five of those rows will be filled and the remainder will contain NA.

So now we have run the first step of the Fama–MacBeth methodology – we have estimated all of the betas, also known as the factor exposures. The slope parameter estimates for the regression of a given portfolio will show how sensitive the returns on that portfolio are to the factors and the intercepts will be Jensen's alpha estimates. These intercept estimates in BETA_C should be comparable to those in the second panel of Table 6 in Gregory *et al.* – their column headed '*Simple 4F*'. Since the parameter estimates in all of their tables are expressed as percentages, we need to multiply all of the figures given from the EViews output by 100 to make them on the same scale. If the 4-factor model is a good one, we should find that all of these alphas are statistically insignificant. We could test this individually if we wished by adding an additional line of code in the loop to save the *t*-ratios in the regressions (something like BETA_T_C(!J)=@TSTATS(2) should do it). It would also be possible to test the joint null hypothesis that all of the alphas are jointly zero using a test developed by Gibbons, Ross and Shanken (1989) – the GRS test, but this is beyond the scope of this book.

The second stage of Fama–MacBeth is to run a separate cross-sectional regression for each point in time. An easy way to do this is to, effectively, rearrange the data so that each column (while still in a time series workfile) is a set of cross-sectional data. So the loop over K takes the observations in the twenty-five portfolios and arranges them cross-sectionally. Thus TIME1 will contain twenty-five data points (one for each portfolio) – all the observations for the first month, October 1980; TIME2 will contain all twenty-five observations for the portfolios in the second month, November 1980; . . . ; TIME387 will contain all twenty-five portfolio observations for December 2012.

We are now in a position to run the second-stage cross-sectional regressions corresponding to equation (14.35) above – note that this runs from 1 to NOBS, which was defined as Gregory *et al.*'s sample to run to December 2010 and not all

Table 14.4 Results from Fama–MacBeth procedure using EViews

Parameter	Estimate	t-ratio
Cons	0.34	0.89
λ_M	0.21	0.62
λ_S	0.08	0.50
λ_V	0.42	2.23
λ_U	0.32	0.50

the data available up to December 2012. Again, it is more efficient to run these in a loop (since there will be 363 of them!) rather than individually. The Z index will loop over each of the months to produce a set of parameter estimates (lambdas) for each one, each time running a regression on the corresponding parameter estimates from the first stage.

Thus the first regression will be of TIME1 on a constant, BETA_RMRF, BETA_UMD, BETA_HML, and BETA_SMB with the estimates being put in new series as before. LAMBDA_C will contain all of the intercepts from the second stage regressions, LAMBDA_RMRF will contain all of the parameter estimates on the market risk premium betas and so on. We also collect the R^2 for each regression as it is of interest to examine the cross-sectional average.

The final stage is to estimate the averages and standard errors of these estimates using something equivalent to equations (14.29) and (14.30) respectively for each parameter. The mean is calculated simply using the @MEAN object, and the standard deviation is calculated using @STDEV. So LAMBDA_C_MEAN will contain the mean of the cross-sectional intercept estimates, and the corresponding t-ratio will be in LAMBDA_C_TRATIO and so on.

Once the program is run, we can double click on each of these objects to examine the contents. The lambda parameter estimates should be comparable with the results in the column headed '*Simple 4F Single*' from Panel A of Table 9 in Gregory *et al*. Note that they use γ to denote the parameters which have been called λ in this text. The parameter estimates obtained from this simulation and their corresponding t-ratios are given in table 14.4. Note that the latter do not use the Shanken correction as Gregory *et al*. do. These parameter estimates are the prices of risk for each of the factors (again, the coefficients from EViews need to be multiplied by 100 to turn them into percentages), and interestingly only the price of risk for value is significantly different from zero. While Gregory *et al*. additionally conduct a range of closely related but more sophisticated tests, their conclusion that further research is required to discover more convincing asset pricing model for the UK is identical to this one using the standard approach.

Table 14.5 Suggested structure for a typical dissertation or project

Title page
Abstract or executive summary
Acknowledgements
Table of contents
Section 1: Introduction
Section 2: Literature review
Section 3: Data
Section 4: Methodology
Section 5: Results
Section 6: Conclusions
References
Appendices

14.11 How might the finished project look?

Different projects will of course require different structures, but it is worth outlining at the outset the form that a good project or dissertation will take. Unless there are good reasons for doing otherwise (for example, because of the nature of the subject), it is advisable to follow the format and structure of a full-length article in a scholarly journal. In fact, many journal articles are, at approximately 5,000 words long, roughly the same length as a student research project. A suggested outline for an empirical research project in finance is presented in table 14.5. We shall examine each component in table 14.5 in turn.

The title page

The *title page* is usually not numbered, and will contain only the title of the project, the name of the author, and the name of the department, faculty, or centre in which the research is being undertaken.

The abstract

The *abstract* is usually a short summary of the problem being addressed and of the main results and conclusions of the research. The maximum permissible length of the abstract will vary, but as a general guide, it should not be more than 300 words in total. The abstract should usually not contain any references or quotations, and should not be unduly technical, even if the subject matter of the project is.

Acknowledgements

The *acknowledgements* page is a list of people whose help you would like to note. For example, it is courteous to thank your instructor or project supervisor (even if he/she was useless and didn't help at all), any agency that gave you the data, friends who read and checked or commented upon the work, etc. It is also 'academic etiquette' to put a disclaimer after the acknowledgements, worded something like 'Responsibility for any remaining errors lies with the author(s) alone'. This also seems appropriate for a dissertation, for it symbolises that the student is completely responsible for the topic chosen, and for the contents and the structure of the project. It is your project, so you cannot blame anyone else, either deliberately or inadvertently, for anything wrong with it! The disclaimer should also remind project authors that it is not valid to take the work of others and to pass it off as one's own. Any ideas taken from other papers should be adequately referenced as such, and any sentences lifted directly from other research should be placed in quotations and attributed to their original author(s).

The table of contents

The *table of contents* should list the sections and sub-sections contained in the report. The section and sub-section headings should reflect accurately and concisely the subject matter that is contained within those sections. It should also list the page number of the first page of each section, including the references and any appendices.

The abstract, acknowledgements and table of contents pages are usually numbered with lower case Roman numerals (e.g. i, ii, iii, iv, etc.), and the introduction then starts on page 1 (reverting back to Arabic numbers), with page numbering being consecutive thereafter for the whole document, including references and any appendices.

The introduction

The *introduction* should give some very general background information on the problem considered, and why it is an important area for research. A good introductory section will also give a description of what is *original* in the study – in other words, how does this study help to advance the literature on this topic or how does it address a new problem, or an old problem in a new way? What are the aims and objectives of the research? If these can be clearly and concisely expressed, it usually demonstrates that the project is well defined. The introduction should be sufficiently non-technical that the intelligent non-specialist should be able to understand what the study is about, and it should finish with an outline of the remainder of the report.

The literature review

Before commencing any empirical work, it is essential to thoroughly review the existing literature, and the relevant articles that are found can be summarised in the *literature review* section. This will not only help to give ideas and put the proposed research in a relevant context, but may also highlight potential problem areas. Conducting a careful review of existing work will ensure that up-to-date techniques are used and that the project is not a direct (even if unintentional) copy of an already existing work.

The literature review should follow the style of an extended literature review in a scholarly journal, and should always be *critical in nature*. It should comment on the relevance, value, advantages and shortcomings of the cited articles. Do not simply provide a list of authors and contributions – the review should be written in continuous prose and not in note form. It is important to demonstrate understanding of the work and to provide a critical assessment – i.e. to point out important weaknesses in existing studies. Being 'critical' is not always easy but is a delicate balance; the tone of the review should remain polite. The review should synthesise existing work into a summary of what is and is not known and should identify trends, gaps and controversies.

Some papers in the literature are *seminal*: they change the way that people have thought about a problem or have had a major influence on policy or practice They might be introducing a new idea or an idea new to that subject area. Reviews can sometimes be organised around such papers and certainly any literature review should cite the seminal works in the field.

The process of writing a literature review can be made much easier if there exists a closely related *survey* or *review* paper. Review papers are published and (usually) high quality and detailed reports on a particular area of research. However, it goes without saying that you should not simply copy the review for several reasons. First, your topic may not match exactly that of the survey paper. Second, there may be more recent studies that are not included in the review paper. Third, you may wish to have a different emphasis and a wider perspective.

An interesting question is whether papers from low ranking journals, poorly written papers, those that are methodologically weak, and so on, be included in the review? This is, again, a difficult balance. In general the answer is probably not, but they should be included if they are directly relevant to your own work, but you should be sure to highlight the weaknesses of the approaches used.

The data

The *data* section should describe the data in detail – the source, the format, the features of the data and any limitations which are relevant for later analysis (for example, are there missing observations? Is the sample period short? Does the sample include large potential structural breaks, e.g. caused by a stock market crash?). If there are a small number of series which are being primarily investigated, it is common to plot the series, noting any interesting features, and to supply summary statistics – such as the mean, variance, skewness, kurtosis, minimum and maximum values of each series, tests for non-stationarity, measures of autocorrelation, etc.

Methodology

'*Methodology*' should describe the estimation technique(s) used to compute estimates of the parameters of the model or models. The models should be outlined and explained, using equations where appropriate. Again, this description should be written *critically*, noting any potential weaknesses in the approach and, if relevant, why more robust or up-to-date techniques were not employed. If the methodology employed does not require detailed descriptions, this section may usefully be combined with the data section.

Results

The *results* will usually be tabulated or graphed, and each table or figure should be described, noting any interesting features – whether expected or unexpected, and in particular, inferences should relate to the original aims and objectives of the research outlined in the introduction. Results should be *discussed and analysed*, not simply presented blandly. Comparisons should also be drawn with the results of similar existing studies if relevant – do your results confirm or contradict those of previous research? Each table or figure should be mentioned explicitly in the text (e.g. 'Results from estimation of equation (11) are presented in Table 4'). Do not include in the project any tables or figures which are not discussed in the text. It is also worth trying to present the results in as interesting and varied a way as possible – for example, including figures and charts as well as just tables.

Conclusions

The *conclusions* section should re-state the original aim of the dissertation and outline the most important results. Any weaknesses of the study as a whole should be highlighted, and finally some suggestions for further research in the area should be presented.

References

A list of *references* should be provided, in alphabetical order by author. Note that a list of *references* (a list of all the papers, books or web pages referred to in the study, irrespective of whether you read them, or found them cited in other studies), as opposed to a bibliography (a list of items that you read, irrespective of whether you referred to them in the study), is usually required.

Although there are many ways to show citations and to list references, one possible style is the following. The citations given in the text can be given as 'Brooks (1999) demonstrated that . . . ' or 'A number of authors have concluded that . . . (see, for example, Brooks, 1999).'

All works cited can be listed in the references section using the following style:

Books

Harvey, A. C. (1993) *Time Series Models*, second edition, Harvester Wheatsheaf, Hemel Hempstead, England

Published articles

Hinich, M. J. (1982) Testing for Gaussianity and Linearity of a Stationary Time Series, *Journal of Time Series Analysis* 3(3), 169–176

Unpublished articles or theses

Bera, A. K. and Jarque, C. M. (1981) An Efficient Large-Sample Test for Normality of Observations and Regression Residuals, *Australian National University Working Papers in Econometrics* 40, Canberra

Appendices

Finally, an *appendix* or *appendices* can be used to improve the structure of the study as a whole when placing a specific item in the text would interrupt the flow of the document. For example, if you want to outline how a particular variable was constructed, or you had to write some computer code to estimate the models, and you think this could be interesting to readers, then it can be placed in an appendix. The appendices should not be used as a dumping ground for irrelevant material, or for padding, and should not be filled with printouts of raw output from computer packages!

14.12 Presentational issues

There is little sense in making the final report longer than it needs to be. Even if you are not in danger of exceeding the word limit, superfluous material will generate no additional credit and may be penalised. Assessors are likely to take into account the presentation of the document, as well as its content. Hence students should ensure that the structure of their report is orderly and logical, that equations are correctly specified, and that there are no spelling or other typographical mistakes, or grammatical errors.

Some students find it hard to know when to stop the investigative part of their work and get to the tidying up stage. Of course, it is always possible to make a piece of work better by working longer on it but there comes a point when further work on the project seems counterproductive because the remaining time is better spent on improving the writing and presentational aspects. It is definitely worth reserving a week at the end of the allocated project time if possible to read the draft paper carefully at least twice. Also, your supervisor or advisor may be willing to read through the draft and to offer comments upon it prior to final submission. If not, maybe friends who have done similar courses can give suggestions. All comments are useful – after all, any that you do not like or agree with can be ignored!